Introduction to Experimental Design and SPSS

Muhammad Waqas

Introduction to Experimental Design and SPSS
Muhammad Waqas

Copyright

This work has been done for the good of humanity and sake of God, not for the rights or reputation of author. Any part of this book **is allowed to copy and distribute without the knowledge of author**; however, must not be manipulated.

Concepts mentioned in this book are purely the knowledge of author and any mistake or error must be conveyed through waqasation@gmail.com or WhatsApp (+92-300-0031923).

Dedicated

TO

Dr. Azeem Ali

(Assistant Professor University of Veterinary and Animal Sciences (UVAS), Lahore, Pakistan)

Dr. Shahid Mehmood

(Associate Professor and Chairman Poultry Production, UVAS, Lahore, Pakistan)

Dr. Muhammad Shabir Shaheen

(Veterinarian and Lecturer UVAS, Lahore, Pakistan)

Foreword

الحمد لله الذي خلق كل شيء وخلق أكمل إنسان على وجه الأرض محمد رسول الله صلى الله عليه وسلم. وجعله مرشدًا للبشرية جمعاء.

All the praises be to God, who created everything and created the most perfect man on the face of Earth, Muhammad (peace be upon him) and made him a guide for all of the mankind.

The moment I enrolled in Experimental Design for postgraduation, I came to know that most of postgraduate students were sacred of this subject. And with a biology background rather mathematics, so was I. However, Dr. Azeem Ali not only taught us to be a true researcher but also proved himself the best teacher. He explained all the concepts like a piece of cake. It would not have been possible for me to get a deep insight into this subject without his master skills. Most of his ideas and concepts have been used in this book but no mistake or error should be attributed to him but to this author alone.

This humble effort is just to make this subject easier for the understanding of new postgraduate students and field researchers. Being a student of science, suggestions and improvements are welcomed, and should be conveyed through author's email (waqasation@gmail.com).

Waqas

Major Contents

Table of Contents

CHAPTER 1: RESEARCH AND EXPERIMENTS

All the knowledge belongs to God.

1.1: STATISTICS AND RESEARCH

Statistics is learning from data**.** It is the science of understanding data and of making decisions in the face of variability and uncertainty. It is concerned with the collection, organization, summarization, and analysis of data, and the drawing of inferences about a body of data when only a part of the data is observed. This analyzed data is then concluded using **statistical inference**, the actual procedure by which we reach a conclusion about a population on the basis of the information contained in a sample drawn from that population.

The **research process** is broadly summarized in Figure 1.1. It begins with an observation that you want to understand, and this observation could be anecdotal or could be based on some data. From initial observation a researcher generates explanations, or theories, of those observations, from which you can make predictions (known as hypothesis). Here's where the data come into the process because to test predictions you need data. First collect some relevant data (and to do that you need to identify things that can be measured) and then you analyze those data. The analysis of the data may support your theory or give you cause to modify the theory. As such, the processes of data collection and analysis and generating theories are intrinsically linked: theories lead to data collection/analysis and data collection/analysis informs theories.
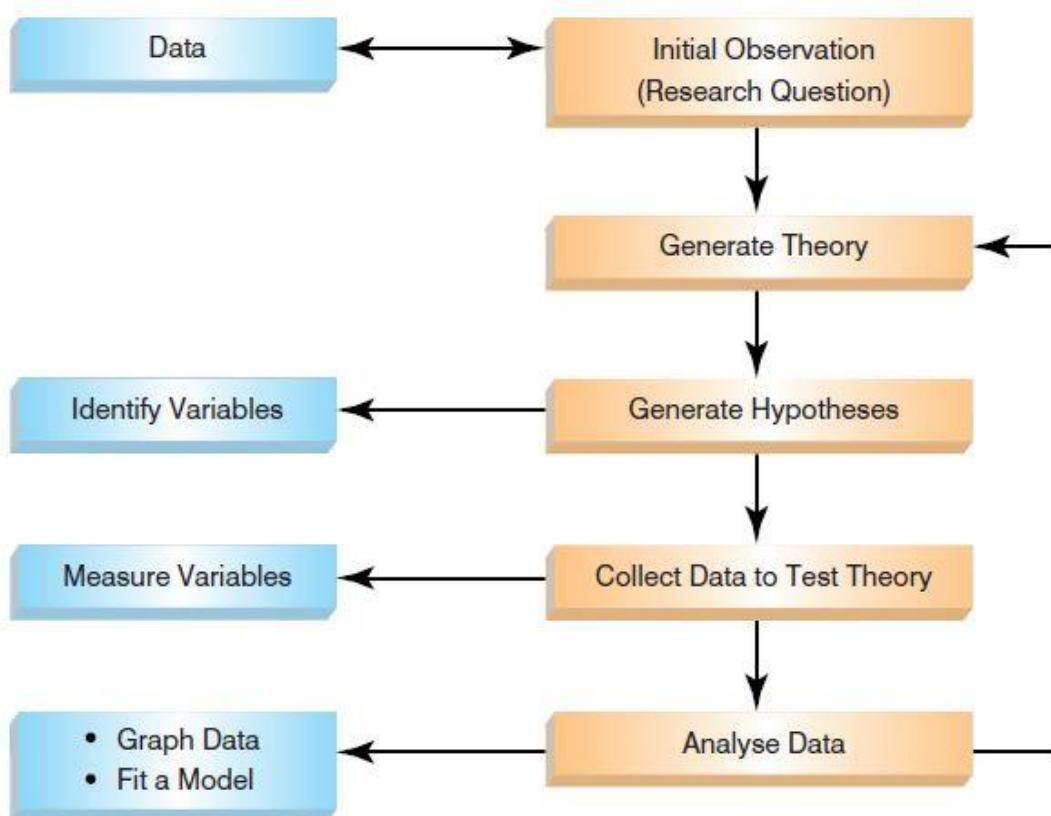
Figure 1.1; research process.

Research in universities is mostly associated with design of experiment (DOE), as the basic theme of the research is 'applied research'. Generally, research can be:

- **Applied research**: can be estimated experimentally using statistics.
- **Pure research**: cannot be estimated on experimental basis.

1.1.1: Pure Research

It is known by several names: academic research; basic research; fundamental research. It is performed without a specific purpose in mind. Rather, it is primarily concerned with generation of the new knowledge. Therefore, the goal of pure research is to produce new

All the knowledge belongs to God.

knowledge and thereby establish principles with which to explain them. It is used to generate and expand theories that describe, explain, or predict a phenomenon of interest to the discipline without regard to its immediate use. It is difficult as it requires pure concepts and new ideas. In other words, it is the formal and systemic process leading to the development of theories.

Pure research deals with the questions that are intellectually challenging to the researcher and may or may not have practical applications at the present time or in the future. That is why, a person wishing to do pure research in any specialized area of science must have studied the concepts and assumptions of that specialization enough to know what has been done and what remains to be done.

1.1.2: Applied Research

The applied research is the scientific investigations conducted to generate knowledge that will directly influence a clinical trial. It refers to those studies that have functional purposes and practical use or application. Applied researchers focus on finding an immediate solution to the exiting problem. They scientifically collect the data to be used in clinical, administrative, or instructional area to find solution to the problem and evaluate practices and identifies needs of the patient.

Applied research is the application of a concept/something which has already been discovered. It has no concern with concept, theory, standards, characteristics, principles and norms. But it is based on statistics which is applied for comparisons in this research like hit and trial. For example, Curie's pure research on radioactive elements led to the applied research of atomic bomb. Remember that theoretical concepts are unchangeable (pure research) unlike applied research.

1.1.3: Social Research

Another type of research, social research, is concerned with the gathering data that can help us answer question about various aspects of society and thus can enable us to understand the society. These questions may pertain to a very specific problems such as how a social welfare caseworker in a city better meet the needs of his/her clients, or how conflict among medical professionals in a particular hospital may be lessened.

Social research is different from above two and is used for social benefits only i.e., if existing studies are copied and applied as it is, for the social benefit, it is social research. But if the existing studies are changed/manipulated and then applied, it will be applied research. There is a wrong myth that 'social research' is 'applied research'. Similarly, applied research is taken as 'pure research' i.e., Newton's second law is said to be applied research but in fact it is related to 'pure research', because it is proved theoretically alone and does not need statistics for its representation. It was, actually, the 'pure' concept of Newton.

1.2: EXPERIMENTS

Experiments are a special type of research study in which observations are made after specific manipulations of conditions have been carried out. They provide the foundation for scientific research and are for finding unknown effects of known factors. It is deliberate observation under conditions deliberately arranged by the observer and '*is a systematic process or series of activities which lead to collection of information on certain aspects to reply to the objectives that the researcher has already in mind*'.

Experiments are only experience carefully planned in advance, and designed to form a secure basis of new knowledge. They enable a researcher to test a hypothesized relationship between an **independent variable** (*suspected casual event that is under investigation*: it is manipulated by the researcher) and **dependent variable** (used to access the effects of manipulating independent variable).

All the knowledge belongs to God.

1.2.1: Types of Experiments

Experiments may be deterministic experiment/approach and probabilistic experiment/approach.

Deterministic experiment *helps predicting the outcome or where the outcome is always fixed/unique.* It is related to 'pure research' and is only followed by the 'pure researchers' as there is only one outcome.

It is based solely on theory and no statistics is involved e.g., O_2 and H_2 always makes H_2O when reacted at certain conditions. Similarly, if a body falls under the gravitational pull, it will always come down as per Newton's law of gravitation. There is no other outcome. But sometimes, same design can give different results due to methodology of process. Because any process contains factors which may be unknown. For example, tossing of coin doesn't give same results (50% head and 50% tail) all the time. Because there may be **deterministic error** in the coin composition or shape just like in the measurements by Vernier Caliper where upon measuring again and again produces different results.

Probability is a mathematical construction that determines the likelihood of occurrence or nonoccurrence of events that are subjected to chance factor. *In **probabilistic approach**, outcome cannot be predicted or in which at least two outcomes are possible.* It is always related to 'applied research' and is so known as **non-deterministic experiment**.

In statistics, only probabilistic approach works. Because if there is no other outcome than one, then there will be no statistics involved. For example, if all the broilers after eating 3.5kg feed in a house produce exactly the same weight (say 2kg) then it will be deterministic approach and no statistics will be involved. If weight varies, then it will be probabilistic approach and statistics must be involved in it. Similarly, the weight of newborn baby and or the tossing of coin cannot be determined before its happening, these are examples of probabilistic approach.

1.3: EXPERIMENTAL STRUCTURE

There are some technical terms which must be understood properly for better understanding of design of experiment (DOE).

1.3.1: Independent and Dependent Variables

An independent variable can be defined as the one that can be manipulated by a researcher. For example, in planning a research experiment to see the effect of different 'intensities of exercise' on the 'performance', exercise intensity is an independent variable because the researcher is free to manipulate it.

A variable is said to be dependent if it changes as a result of the change in the independent variable. In the above example, 'performance' is a dependent variable because it is affected by the change in 'exercise intensity'. In fact, dependent variable can be taken as the variable of interest. In creating the graph, the dependent variable is taken along the Y-axis, whereas the independent variable is plotted on the X-axis.

1.3.2: Extraneous Variable

These are variables *that are not of primary interest in a particular study but could influence the dependent variable* and are also known as **nuisance variables** or, in some designs, **covariates**. Environmental factors (e.g., temperature or distractions), time of day, and characteristics of the experimenter, teacher, or therapist are some possible extraneous variables that need to be controlled.

1.3.3: Treatments and Factors

There is a minor difference between treatment and factor and sometimes taken as synonymously. *Treatments are the set of conditions under study or different procedures we want to compare.* These are the experimental condition whose effect is to be measured and compared e.g., these could be different kinds or amounts of feeds in poultry or fish breeding.

All the knowledge belongs to God.

Factor, on the other hand, is a set of treatments of a single type and is a concept used to denote a group of treatments. For example, different breeds of cattle, different diets, different varieties, different doses of nitrogen, different methods of irrigation, etc. may form different factors in factorial experiments.

Factors combine to form treatments e.g., the baking treatment for a cake involves a given time at a given temperature. The treatment is the combination of time and temperature, but we can vary the time and temperature separately. Thus, we speak of a time factor and a temperature factor. The treatments, or '**treatment combinations**,' then can consist of all combinations of factors and factor levels or just a subset of the possible combinations. For example, if time and temperature each have 3 levels (low, medium, and high), then the full set of treatments that can be constructed from these two factors are the 9 possible combinations of time and temperature levels. *When these treatment combinations are tabulated or entered into a spreadsheet, the suite of experiments is called an **experimental matrix***. Remember that to change the treatment factor in small animals (chicken or fish) you'll have to replace the animals due to extra-ordinary changes in parameters in short time; however, in large animals (say buffalo), these are not replaced but written as **phase-out process** of life.

1.3.4: Levels of Factor

Individual settings for each factor or the subsets produced by the factor are called levels of the factor. When someone refer to 'factor levels' it means different 'expressions of that factor'. For example, different settings (200°C and 300°C) for the factor 'temperature', or different therapies (radiation and chemotherapy) for the factor 'cancer treatment'. These settings and therapies are levels. Similarly, different breeds of cattle and different types of feed are levels of the factors breed and diet, respectively. Hence, these levels are said to be different components of a factor. Both the factors and their levels may be quantitative (e.g., doses of feed) as well as qualitative (different types of diets, different varieties of grains) in nature.

1.3.5: Experimental and Measurement Units

Experimental units (EU) are the things to which we apply the treatments or the entities that receive an independent application(s) of one of the experiment's treatment(s). In other words, the smallest part/unit of the experimental area/subjects (e.g., plots in field experiment, pots, test tubes, etc.) in which one applies treatments and from which observations/responses are recorded is EU. It is assigned to an experimental condition independently of other entities. However, the objects or the units on/in which treatments are applied are known as **subjects**.

EUs should always be homogenous units. These could be strains of broilers, groups of deer receiving different feeds. These should be identical and performed under the same environments/conditions. But these units should be randomly selected and be representative of the population.

Measurement units, also known as **observation unit (OU)** or **sampling units**, are the actual objects on which the response is measured. These may differ from the EUs and so a careful distinction must be made between them. For example, consider the effect of different fertilizers on the nitrogen content of corn plants. Different field plots in this experiment are the EUs, but the OUs might be a subset of the corn plants on the field plot, or a sample of leaves, stalks, and roots from the field plot. Similarly, in an educational study, a class, that is, a collection of students, is the EU as the class as a whole is subjected to a particular teaching method which is the treatment. Observations are made; however, on the individual students in the form of test scores. Then the students are the OUs.

All the knowledge belongs to God.

1.3.6: Parameters and Confounding

A parameter is a number describing a whole population or a measured quantity in a statistical population. Every experiment has some parameters i.e., Chat GPT4 (which works on search and artificial intelligence methodology) has about 400 trillion parameters. However, high number of parameters means greater number of checks, thus high complexity. But this complexity produces more reliable response/forecast.

Confounding occurs when the effect of one factor or treatment cannot be distinguished from that of another factor or treatment. Thus, the two factors or treatments are said to be confounded. Consider planting corn variety A in Islamabad and corn variety B in Murree. In this experiment, we cannot distinguish location effects from variety effects—the 'variety' factor and the 'location' factor are confounded here. Remember that, except in very special circumstances, confounding should be avoided.

1.3.7: Predicator and Response

In research, one possibility is to abandon the terms dependent and independent variable and use the terms predictor and outcome variable. *Those parameters which control the outcomes of that experiment, are known as predictors (thought to predict an outcome variable) and the outcome is called response.*

Response is the outcome that we observe after applying a treatment to an EU. It is the 'measured characteristics' used to evaluate the effect of treatments on EUs, e.g., in experiment of effect of temperature of house on broiler weight gain, temperature will be predicator and weight gain as response. The response/findings may be positive (due to consistent methodology and low variability) and negative and less acceptable (due to less consistency and more variation). For example, a good tea maker in home is more consistent with less variation in methodology of tea making. Remember during interpreting a problem, if the response is nominal in stated problem, then there is no DOE applicable for the problem.

1.3.8: Control and Placebo Groups

Control treatment is a 'standard' treatment that is used as a baseline or basis of comparison for other treatments. This treatment might be the treatment in common use, or it might be a 'null treatment' (no treatment at all). Similarly, a **control group** *is the group that receives no treatment at all*, also known as **placebo group**. Control treatment, in fact, reduces the chance of experimental error and helps identify variation in experiment.

Placebo is a null treatment that is used when the act of applying a treatment—any treatment—has an effect. It is the 'psychological effect'. It is actually an inert substance or neutral stimulus that is administered to subjects as if it was the actual treatment condition. That is why any change in the dependent variable attributable to receiving a placebo is called **placebo effect**.

1.3.9: Blinding

In experiments on humans, particularly those that involve the use of placebos, blinding is often used. It means *'the treatment assignment is kept secret from the experimental subject'.* Blinding also occurs when the evaluators of a response do not know which treatment was given to which unit. The purpose of blinding the subject/evaluator is to minimize the extent to which his or her expectations influence the results of the experiment. *An experiment in which both the subjects and the persons making the evaluations of the response are blinded is called a **double-blind experiment**.* Following is an example of a double-blind randomized controlled crossover trial.

The purpose of an investigation was to evaluate the analgesic effectiveness of a daily dose of oral methadone in patients with chronic neuropathic pain syndromes. The researchers used a visual analogue scale (0–100 mm, higher number indicates higher pain) ratings for maximum pain intensity over the course of the day. Each subject took either 20mg methadone or a placebo each day for 5 days. Subjects and evaluator did not know which treatment they

All the knowledge belongs to God.

were taking. The response was mean maximum pain intensity scores for the 5 days on methadone and the 5 days on placebo.

1.3.10: Errors

Error is the variation in response and may be of two types:

- Experimental error: may be due to various sources of variations among different EUs may be ascribed due to:
    - A systematic part (assignable part)
    - A nonsystematic part (non-assignable part)
- Statistical error

Systematic variation/error part is consisting of that part of the variations caused due to known sources of variations like differences in treatments, blocks, etc. It is an error that it arises from variation that is uncontrolled and generally unavoidable. These errors occur while collection of data and may be due to **sampling** and **non-sampling errors**. Statistical error should be minimized to every extent.

But the part of the '*variation which cannot be assigned to specific reasons or causes, i.e., the non-systematic part, is termed as the* **experimental error'**,* i.e., the individual runs differ in the same experiments, so there is fluctuation, or noise, in the observed experiment. This noise is usually called experimental error or simply error. While designing and planning of any experiment, a researcher always intends to minimize the experimental error. Experimental error is the random variation present in all experimental results i.e., different EUs will give different responses to the same treatment, and it is often true that applying the same treatment over and over again to the same unit will result in different responses in different trials. Remember that experimental error does not refer to conducting the wrong experiment. It is associated with inherit variability of EUs, methodologies or instruments. Such error is a reflection of the fact that EUs or instruments are not alike, that is, cannot be replicated exactly.

Contributing to experimental error is also our failure to replicate a treatment exactly, that is, instead of administering 15 ppm of a certain substance, as called for in the protocol, we administer to some units 14 ppm or 16 ppm and so on. We refer to this component of the experimental error as **treatment error**. Two main sources of experimental error are:

- **Inherent variability** in the experimental materials in which the treatments are applied.
- **Lack of uniformity** in the physical conduct of the experiment or in other words, failure to standardize the experimental techniques.

1.3.11: Degree of Freedom

In statistical terms the degrees of freedom relate to the number of observations that are free to vary. If we take a sample of four observations from a population, then these four scores are free to vary in any way (they can be any value). However, if we then use this sample of four observations to calculate the standard deviation of the population, we have to use the mean of the sample as an estimate of the population's mean. Thus, we hold one parameter constant.

Suppose that the mean of the sample was 10; then we assume that the population mean is 10 also and we keep this value constant. With this parameter fixed, can all four scores from our sample vary? The answer is no, because to keep the mean constant only three values are free to vary. For example, if the values in the sample were 8, 9, 11, 12 (mean = 10) and we changed three of these values to 7, 15 and 8, then the final value must be 10 to keep the mean constant. Therefore, if we hold one parameter constant then the degrees of freedom must be one less than the sample size. This fact explains

All the knowledge belongs to God.

why when we use a sample to estimate the standard deviation of a population, we have to divide the sums of squares by 'N-1' rather than 'N' alone.

All the knowledge belongs to God.

CHAPTER 2: DATA AND ASSUMPTIONS

2.1: DATA

The raw material of statistics is data and is usually written as *'a collection of measurements or observations'*. Data are numbers which contain information, and the purpose of statistics is to investigate and evaluate the nature and meaning of this information. That is why data should be reliable and verifiable.

Data can be collected from the existing sources, or these can be generated from experiments conducted for the purpose adopting.

- Complete enumeration
- Sampling technique

In **complete enumeration** technique (census), data are collected from each and every individual unit of the targeted population. But in many situations, it may not be possible or feasible (because of time, financial, accessibility, or other constraints) to study each and every individual element of interest, resulting in the selection of a representative part (sample) of the study objects (population) using appropriate **sampling technique**.

2.1.1: Between and Within-subject Factors

When we collect data in an experiment, we can choose between two methods of data collection. The first is to manipulate the independent variable using different participants. In this method different groups of people take part in each experimental condition (a **between-groups, between-subjects or independent design**). The second method is to manipulate the independent variable using the same participants (a **within-subject or repeated measures design**). Simplistically, this method means that we give a group of students, positive reinforcement for a few weeks and test their statistical abilities and then begin to give this same group negative reinforcement for a few weeks before testing them again, and then finally giving them no reinforcement and testing them for a third time. These two factors are very useful in postgraduate studies and must not be mixed.

2.1.2: Types of Data

Data depending upon the source of its collection/generation or collation may be:

1) Primary data
2) Secondary data

Primary data is generated by the investigator/experimenter through a well-planned program for specific purpose and may be obtained through survey or conducting field experiments etc. Thus, primary data is generated by the user for specific purpose. For example, data collected on egg-laying capacity of particular poultry breed under particular management practice by researcher himself. But the source of collecting primary data may be:

1) **Instrument**: it includes a questionnaire, survey form, Google ads e.g., just asking someone which brand of tea do you like, without doing the actual experiment. If the data needed to answer a question are not available from routinely kept records, the logical source may be a survey. Suppose, for example, that the administrator of a clinic wishes to obtain information regarding the mode of transportation used by patients to visit the clinic. If admission forms do not contain a question on mode of transportation, we may conduct a survey among patients to obtain this information.
2) **Experiment**: it involves actually doing something like presenting 3 cups of 3 different brands of tea to each participant for their liking or disliking. But it is always a probabilistic experiment.

All the knowledge belongs to God.

3) **Instrument with experiment**: providing a questionnaire along with the actual experiment i.e., asking since how many years have you been using your favorite tea brand after actually presenting the tea to the participant.

The last one is the most effective among all options of primary data collection.

Secondary data is the data used by the experimenter or user, which are collated by someone else (from other sources). For example, weather data are recorded by the department of meteorology (primary data), one of their primary objectives or mandates; but many agencies like the airport authority, agriculture department, disaster management department, and the experimenters/researchers in biological sciences use these weather data collating from the meteorology department in order to explain more meaningful way the phenomenon under their considerations (making it secondary data).

Why do research papers get rejected? Research papers are mostly rejected based on flawed data methodology i.e., the collection of data should be justified according to the project requirement. Similarly, if there is a project of one broiler flock for 42 days. There must be certain parameters used in that research (data collection) which could justify the budget and time of the project. If a number of papers site that flock reference without participating in it, it can be judged by the budget/time-period calculations of that research flock whether it was suspicious or genuine? The other reason for postgraduate research failure in the field is the **OFAT design** (mentioned later).

2.2: STEPS TO SOLVE A STATISTICAL PROBLEM

In probabilistic experimental design, steps to solve a problem include:

1) Collection of data
2) Organization of data
3) Forecasting

Statistical sources or techniques are introduced into the above 3 steps for organization of large sets of data for more efficiency e.g. statistical software: SAS; GenStat; MINTAB; SPSS; SYSTAT; INSTAT; Agrobase; MSTATC etc.

How does Google Maps predict route and travelling? It actually collects data from all the sources present on a route (who are connected), organize into different sections on the basis of speed i.e., it categorizes them into pedestrians; on cycle; on motorbike; and on cars. Then it forecast speed and time. Similarly, it predicts the traffic jams on the routes by collecting data from passengers already moving on the very route.

How to start a business of burger? Well! one must think about the survival of business first, at least for 100 years. Above mentioned steps should be used for starting a business:

1) Collect data of market or place where to start the business; check already existing burger businesses; find out what is the missing or disturbing part?
2) Then organize your data what you are looking into it or what you need to focus? Organization will give you a pattern.
3) Finally forecast the outcome.

Can a superstore be successful in a small town or countryside? Yes. Because it will provide everything under one roof and maintain its stock due to data preset in its system using real time statistical methodology. It will use above mentioned 3 steps for order placement and stock maintenance. On the other hand, small retailers will be unable to maintain stocks in peak times of year and will depend on the personal memories to stockpile their stuffs. There will be a major design difference in both of these situations.

All the knowledge belongs to God.

2.3: DATA MEASUREMENT LEVEL

Measurement is the assignment of numbers to different characteristics according to rules. The way these numbers are assigned differs across variables, and these differences can affect which statistical tests are appropriate for the data. These differences in the way numbers are assigned can be classified according to the "levels of measurement" that they represent. In other words, *the relationship between what is being measured and the numbers that represent what is being measured is known as the **level of measurement***.

Depending on the level of measurement of a variable, the data can mean different things. For example, the number 2 might indicate a score of two; it might indicate that the subject was a Catholic; or it might indicate that the subject was ranked second in the class. To help understand these differences, types, or levels of variables, have been identified.

Data measurement level (DML) helps what parameters to look for exactly i.e. what gender or what number to look for in research. A variable has one of four different levels of measurement, from lowest to highest, are:

1) Nominal
2) Ordinal
3) Interval
4) Ratio

SPSS (statistical package for social science) uses three terms (nominal, ordinal, and scale) for the levels or types of measurement. Scale is used for interval and ratio.

2.3.1: Variables

A variable is a characteristic of a person or a thing that can be assigned a number or a category. In other words, observable or measurable properties of the observation unit which can take different values is variable. For example, blood type (A, B, AB and O) and age are two variables we might measure on a person. Similarly, a variable is a phenomenon that changes from time to time, place to place, and individual to individual. It can be numeric or attribute.

Variable is a characteristic of the participants or situation for a given study that has several values in that study. A variable must be able to vary or have different values or levels in the study. For example, gender is a variable because it has two levels, female or male. Age is a variable that has a large number of values. Variables are actually the parameters. Remember that variability in experiment should be minimum or lowest.

2.3.1.1 Variable versus Factor

Variable and factor are two different things, variable has a complete set of information with it and factor does not. Factor is a multi-dimension thing. *A mixture of variable is a factor*. The list of questions never ends in the case of factor. For example, growth and satisfaction are factors, but data of height of class students is variable as it has no question/query after the data has been presented. That is why factors always keep research dimensions open- has more room for information. Remember that applied researcher play with factor and not with the variable.

2.3.1.2 Random Variable

Whenever we determine the height, weight, or age of an individual, the result is frequently referred to as a value of the respective variable. When the values obtained arise as a result of chance factors, so that they cannot be exactly predicted in advance, the variable is called a random variable. An example of a random variable is adult height. When a child is born, we cannot predict exactly his or her height at maturity.

All the knowledge belongs to God.

2.4: TYPES OF VARIABLES

Treatments can be structured according to multiple factors or variables. Treatment variables/factors can be either:

- Qualitative variable: can be divided into
  - o Nominal variable
  - o Ordinal variable
- Quantitative variable: can be divided into
  - o Discrete variable
  - o Continuous variable: are further divided into
    - ▪ Interval
    - ▪ Ratio

2.4.1: Qualitative/Categorical Variables

Qualitative data is a categorical measurement and is expressed not in terms of numbers, rather by means of a natural language description. It is often known as **categorical data**. Qualitative characters are also known as **attributes**. *The claim which is given after observation or characters which cannot be quantified exactly but can be categorized/grouped/ranked are known as qualitative characters*. Religion (viz., Hindu, Muslim, Christian etc.), gender (male/female, boys/girls), color (viz., blue, red, green, etc.), and complexion (bad, good, fair, etc.) are the examples of qualitative character. Qualitative variables may be of two types: nominal variables; ordinal variables.

Nominal and ordinal variables are both discrete. The term is also sometimes used for nominal and/or ordinal data, categorical reflecting the various categories that the values of the variable represent.

2.4.1.1 Nominal Variable

As the name implies it consists of 'naming' observations or classifying them into various mutually exclusive and collectively exhaustive categories. The values of a nominal variable allow us to classify groups that have no intrinsic order. It is the **first level in DML** and is always observable. *It has **no preferences at all** in research or has equal preferences in case of quality*. These are just named categories. In this level of measurement, the numbers in the variable are used only to classify the data. However, words, letters, and alpha-numeric symbols can also be used. This scale has certain characteristics but doesn't have any form of numerical meaning. The data can be placed into categories but can't be multiplied, divided, added or subtracted from one another. It's also not possible to measure the difference between data points.

The values we use to designate the categories of such a nominal variable are completely arbitrary. Variables measured on this scale are known as **categorical variables**. Categorical variables result from a selection of categories. Examples might be response (agree, disagree), sports specialization, race, religion, etc. If in a class 30 subjects are male and 20 are female, no gradation is possible. In other words, 30 do not indicate that the males are better than the female in some sense. Similarly, suppose there are data about people belonging to 3 different gender categories. In this case, the person belonging to the female gender could be classified as F, the person belonging to the male gender could be classified as M, and transgendered classified as T. This type of assigning classification is nominal level of measurement.

All the knowledge belongs to God.

Never use nominal variables in your research as it is always used in case of contradiction. And there is no contradiction in nominal factor. Remember that, nominal is coded in 0 and 1 form, but interval is coded in 1 and 2 form.

2.4.1.2   Ordinal Variable

The ordinal level of measurement indicates an ordering of the measurements. Here the categories are ordered. *It allows you to rank the values. It has preference for at least one qualitative variable*, such as for education level: elementary is less than high school, which in turn is less than university. Categorical variables that assess performance (good, average, poor, etc.) are ordinal variables. Similarly, the variables that measure attitude (strongly agree, agree, undecided, disagree, and strongly disagree) are also ordinal variables. On the basis of the order of these variables, we may not know the magnitude of the measured phenomenon of an individual, but we can always grade them. For instance, if A's playing ability in soccer is good and B's is average, we can always conclude that the A is better than B, but how much is not known. Moreover, the distance between the ordered categories is also not same and measurable.

The values of variables reflect order but not magnitude. Cases with higher values on an ordinal variable have more of the construct captured by the variable than those with lower values, but the specific values of the numbers are arbitrary. It can be further elaborated as, suppose a student scores the highest grade of 100 in the class. In this case, he would be assigned the first rank. Then, another classmate scores the second highest grade of 92; she would be assigned the second rank. A third student scores 81 and he would be assigned the third rank, and so on. While each value is ranked, there's no information that specifies what differentiates the categories from each other. These values can't be added to or subtracted from.

2.4.2: Quantitative Variable

Quantitative (also known as numerical) data is a numerical measurement expressed in terms of numbers. There must be an instrument or tool for measurement for this type of variables. *The characters which can be quantified and measured are known as quantitative characters,* i.e. height, weight, age, income, expenditure, production, disease severity, percent disease index, etc. It is not necessary that all numbers are continuous and measurable. For instance, the roll number is a number, but not something that one can add or subtract. Quantitative data are always associated with a scale measure. Quantitative variables can be divided into:

Discrete variables: A discrete variable is characterized by gaps or interruptions in the values that it can assume. These gaps or interruptions indicate the absence of values between particular values that the variable can assume. Some examples illustrate the point. The number of daily admissions to a general hospital is a discrete random variable since the number of admissions each day must be represented by a whole number, such as 0, 1, 2, or 3. The number of admissions on a given day cannot be a number such as 1.5, 2.997, or 3.333. The number of decayed, missing, or filled teeth per child in an elementary school is another example of a discrete variable.

Continuous variables: A continuous random variable does not possess the gaps or interruptions characteristic of a discrete random variable: it include the various measurements that can be made on individuals such as height, weight, and skull circumference. Continuous variables can be measured on two different types of scales:

- Interval variable
- Ratio variable

Interval and ratio levels of measurement are sometimes called **continuous** or **scale**.

All the knowledge belongs to God.

2.4.2.1   Interval Variable

The interval scale is a quantitative measure. It is a set of prefixed criteria/target. The interval level of measurement not only classifies and orders the measurements, but it also specifies that the distances between each interval on the scale are equivalent along the scale from low interval to high interval. Thus, this type of data shows both the order of the variables and the exact differences between the variables. They can be added to or subtracted from each other, but not multiplied or divided. This scale is also characterized by the fact that the number zero is an existing variable. In the ordinal scale, zero means that the data does not exist. In the interval scale, zero has meaning – for example, if you measure degrees, zero has a temperature.

It also has an equidistant measure. But the **doubling principle** breaks down in this scale. The 4 marks given to an individual for his creativity do not explain that his nature is twice as good as the person with 2 marks. This is so because on this scale zero cannot be exactly located. Thus, variables measured on an interval scale have values in which differences are uniform, but ratios are not.

The values on these variables reflect both order and magnitude. Any two values that are interval separated by the same numeric distance are separated by the same amount of the construct captured by the variable. Examples include earnings measured in dollars or distance measured in miles. Two people who earn $60,000 and $50,000 are separated by as much on their salaries as two people who earn $20,000 and $10,000. And, commuters who travel 35 versus 40 miles to work each day are separated by the same distance as those who travel 5 versus 10 miles to work each day.

It is possible to put these variables in an order that is low, medium and high. For example, low, middle and high class income groups based on the previous set target of some income i.e. low class with income of <50k rupees per month, middle class with income from 50-200k rupees and high income class >200k rupees per month. Similarly, egg grading on weight basis into different categories is example of interval variable i.e. jumbo (>70g), XL egg (63-70g); large egg (57-63g); small egg (45-57g) and peeve egg (<45g). CGPA is also an example of interval variable. A popular example of this level of measurement is temperature in centigrade, where, for example, the distance between 94ºC and 96ºC is the same as the distance between 100ºC and 102ºC.

Note that interval variable and categorical variable are sometimes easily interchangeable i.e. if the ranking is based on observation then it will be categorical and if it is based on some measurement tool/instrument, then it will be interval variable.

2.4.2.2   Ratio

The highest level of measurement is the ratio scale. Continuous variables can take on an infinite set of values, typically with decimals representing any value between two integers. This scale is characterized by the fact that equality of ratios as well as equality of intervals may be determined. The data on ratio scale has a meaningful zero value and has an equidistant measure (i.e., the difference between 30 and 40 is the same as the difference between 60 and 70). For example, 60 marks obtained in a test is twice that of 30. This is so because zero exists in the ratio scale. Height is another ratio scale quantitative measure. Observations that are counted or measured are ratio data (e.g., number of goals, runs, height, and weight).

Ratio scales of measurement include properties from all 4 scales of measurement. The data is defined by an identity (nominal), can be classified in order (ordinal), contains

intervals (interval) and can be broken down into exact value (ratio). In this level of measurement, the observations, in addition to having equal intervals, can have a value of zero as well. The zero in the scale makes this type of measurement unlike the other types of measurement, although the properties are similar to that of the interval level of measurement. Ratio scales differ from interval scales in that this scale has a **'true zero'.** The number zero means that the data has no value point. In the ratio level of measurement, the divisions between the points on the scale have an equivalent distance between them.

In simple words, if the ranking or categorization of something is based on per unit basis then it will be ratio factor. For example, take watermelon sale on a superstore. If the watermelons are divided into 2 categories say <400g and >400g, then it will be interval variable. If the seller sell watermelon on per kg base price say 2 rupees per gram, then it will be ratio variable. Similarly, marks percentage is a ratio variable. Remember that ratio is the most precise in measurement and has the most information.

2.4.3: Summary of DML

The researcher should note that among these levels of measurement, the nominal level is simply used to classify data and ordinal is used to rank the data, whereas the levels of measurement described by the interval level and ratio levels are much more exact. According to the amount of information, these DMLs can be organized as:

$$Ratio > Interval > Ordinal > Nominal$$

Note that we can go from ratio variable towards nominal but cannot go backward from nominal to ratio variable. To summarize, nominal scales are used to label or describe values. Ordinal scales are used to provide information about the specific order of the data points, mostly seen in the use of satisfaction surveys. The interval scale is used to understand the order and differences between them. The ratio scales give more information about identity, order and difference, plus a breakdown of the numerical detail within each data point.

2.5: ASSUMPTIONS

Assumption is different from a **constraint**, which is the limitation/restriction imposed during research/study trial. There are two types of constraints:

1) Study constraints: like budget availability
2) Available constraints: like instrument available to researcher in lab.

Every statistical test has assumptions. Statistical assumptions are much like the directions for appropriate use of a product found in an owner's manual. Assumptions explain when it is and isn't reasonable to perform a specific statistical test. When the t-test was developed, for example, the person who developed it needed to make certain assumptions about the distribution of scores, etc., in order to be able to calculate the statistic accurately. If these assumptions are not met, the value that SPSS calculates, which tells the researcher whether the results are statistically significant, will not be completely accurate and may even lead the researcher to draw the wrong conclusion about the results.

We analyze experimental results by comparing the average responses in different treatment groups using an overall test based on ANOVA or more focused procedures. All these procedures are based on the assumption. If we apply the model-based methods in situations where the model assumptions do not hold, the inferences we obtain may be misleading. The three basic assumptions we need to check are that the errors are:

1) Independent (independence)
2) have constant variance or homogenous (homogeneity)

All the knowledge belongs to God.

3) normally distributed (normality)

The fourth assumption, interval data is sometimes taken as 4th assumption. If the data is nominal, parametric tests cannot be applied. Similarly, **linearity** is sometimes mentioned as assumption, (explained later in details). These three assumptions tell us if there is any impact on outcome due to known/unknown causes/factors. All three assumptions are necessary for the applied researcher and they reduce the human errors.

2.5.1: Independence

This assumption ensures that the data is independent of any biasness or drawn from an independent population. Observation on one unit should not be influenced by observation on another or the behavior of one participant does not influence the behavior of another. For example, if you know one subject's value on a variable (e.g., competence), then this should not help you to guess the value of that variable for any other particular participant. Sometimes, this assumption is violated because one's procedures for sampling participants create systematic bias. For example, "snowball sampling," in which participants recommend other participants for the study, is likely to lead to non-independence of observations because participants are likely to recommend people who are similar to themselves. Obviously, members of the same family, or the same person measured on more than one occasion, do not comprise independent observations.

As an another example, imagine two people, Paul and Julie, were participants in an experiment where they had to indicate whether they remembered having seen particular photos earlier on in the experiment. If Paul and Julie were to confer about whether they'd seen certain pictures then their answers would not be independent: Julie's response to a given question would depend on Paul's answer, and this would violate the assumption of independence. If Paul and Julie were unable to confer (if they were locked in different rooms) then their responses should be independent (unless they're telepathic): Paul's responses should not be influenced by Julie's.

In case of violation (dependence), this means that our estimates of standard errors for treatment means and contrasts are biased (whether too large or small depends on the pattern of dependence).

Serial dependence or autocorrelation is one of the more common ways that independence can fail. Serial dependence arises when results close in time tend to be too similar (**positive dependence**) or too dissimilar (**negative dependence**). Positive dependence is far more common. Serial dependence could result from a "drift" in the measuring instruments, a change in skill of the experimenter, changing environmental conditions, and so on. If there is no idea of time order for the units, then there can be no serial dependence.

Spatial association, another common form of dependence, arises when units are distributed in space and neighboring units have responses more similar than distant units. For example, spatial association might occur in an agronomy experiment when neighboring plots tend to have similar fertility, but distant plots could have differing fertilities.

There are no simple methods for dealing with dependence (accommodating dependence) in data; however, independence can be achieved by randomization.

Remember that in repeated-measures designs (in which participants are measured in more than one experimental condition), we expect scores in the experimental conditions to be non-independent for a given participant, but behavior between different participants should be independent.

All the knowledge belongs to God.

2.5.2: Homogeneity or Constant Variance

Homogeneity is the same as **constant variance**. When means from different populations are to be compared, one important assumption is that the populations from which the means are obtained roughly have similar variance. If the variances of different populations are different, it is not possible to pool variance and perform significance test. This is an important assumption to be checked when working with multi-location and several-year trial, and also when a combined analysis of variance is required.

Homogeneity of EUs is achieved at all costs: except factors, all others things should be fixed. For example, if protein is factor in feed, then all other things in feed i.e. energy, fiber, fats etc. should be same in feeds for all replicates. Sometimes, the number of EUs provided/available are not sufficient to be equally replicated, then we will have to make them equal using principle of homogeneity.

In designs in which you test several groups of participants this homogeneity assumption means that each of these samples comes from populations with the same variance. In correlational designs, this assumption means that the variance of one variable should be stable at all levels of the other variable.

Homogeneity is also known as **homoskedasticity**. It states that the dispersion among levels of factors should be constant or the variation among treatments should be constant (equal spreadness from constant). In other words, it assumes that the variance of the errors are constant across levels of factors or the proportionate variation in outcome due to variation in factorial treatment. If there comes any deviation in outcome, researcher can check that there is a problem in treatments and what it can be? So, this assumption can help making a 'guess'. Remember that if errors are not constant across levels of factors, this condition is heteroskedasticity.

Remember that at certain point the outcome will either stop changing (remain constant) or start decreasing with the change of factors. For example, if protein is a factor to be studied for weight gain and there are 4 levels of protein say 0%, 5%, 10% and 15%. There will be a certain trend of increase in weight gain with increase of protein level. This trend of weight gain will vary proportionately in each treatment. If there is say 100g gain in 5% protein feed, there may be about 200g in 10% and 300g in 15%. This is known as constant variance. This trend is usually in the knowledge of researcher. But the increase in weight gain will stop at certain point of protein level. In other words, the weight gain for 5% protein should have same spreadness as increased by 10%.

2.5.3: Normality

Normality talks about overall experiment: the normal distribution is symmetric about its center. It means that all outcomes should be consistent or the mean outcome of all the means should be in the center of graphically represented data and should be maximum. In other words, the shape of the graphical representation of means should be bell-shaped. Remember that the chance of getting this shape subjectively is very rare and is not normal. That is why normality is always taken as 'approximate normality'. If the number of samples are large, it will be easy to draw the normality curve/line smoothly and vice versa. So take large number of samples. If the normality is not achieved, then there must be uncertainty in the experiment/treatment or there may be problem in the planning and execution of experiment.

Keep in mind that if normality is assumed, then the researcher should choose parametric version of DOE otherwise non-parametric version of DOE. Parametric experimental design is a design method where features (such as building elements and engineering components) are shaped according to algorithmic processes. The non-parametric designs do not assume the shape of the data and we have to estimate the most suitable form of the model.

All the knowledge belongs to God.

2.5.3.1   Normal Probability Plot

The normal probability plot (NPP), sometimes called a **Rankit Plot**, is a graphical procedure for assessing normality (Figure 2.1). We plot the ordered data on the vertical axis against the ordered normal scores on the horizontal axis. The normal distribution, as a mathematical function, is characterized by two parameters—two quantities in the mathematical expression for the distribution. These parameters determine where the curve is centered and how spread out it is.
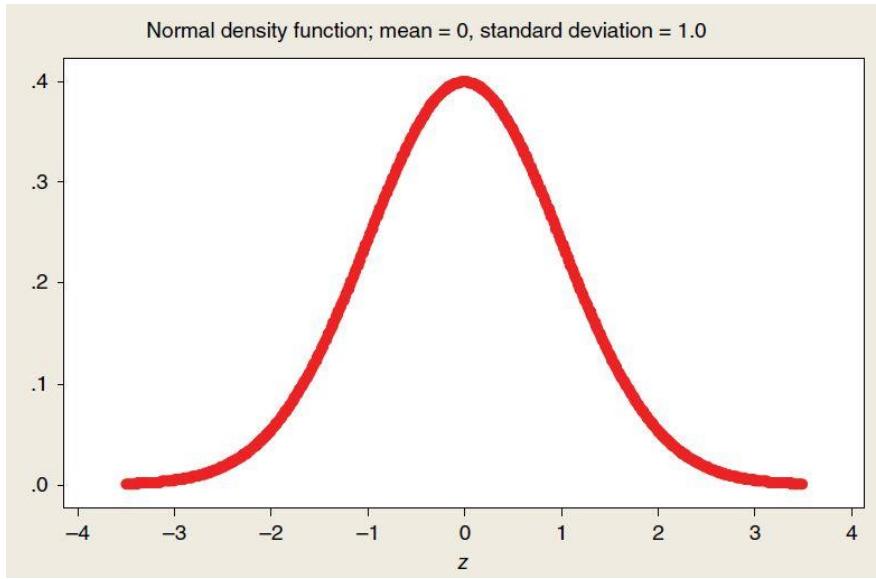
Figure 2.1: Normal probability plot.

2.5.3.2   Outliers

An outlier is a value that falls far from the rest of the values on a variable. These are an extreme form of non-normality. Roughly speaking, an outlier is an observation "different" from the bulk of the data, where different is usually taken to mean far away from or not following the pattern of the bulk of the data. For example, sometimes there are certain animals (say broilers) that either weigh too much or too short of the mean weight at the end of trial, these are the outliers.

An outlier might occur because of a recording error or typographical error when the data are recorded, because of an equipment failure during an experiment, or for many other reasons. Outliers are the most interesting points in a data set. Sometimes outliers tell us about a problem with the experimental protocol (e.g., an equipment failure, a failure of a patient to take his or her medication consistently during a medical trial). At other times an outlier might alert us to the fact that a special circumstance has happened (e.g., an abnormally high or low value on a medical test could indicate the presence of a disease in a patient).

2.5.3.3   Accommodating Non-normality

Non-normality, particularly asymmetry, can sometimes be lessened by transforming the response to a different scale. Skewness transformations to the right is lessened by a square root, logarithm, or other transformation to a power less than one, while skewness to the left is lessened by a square, cube, or other transformation to a power greater than one.

All the knowledge belongs to God.

Individual outliers can affect our analysis. It is often useful to perform the analysis both with the full data set and with outliers excluded. If your conclusions change when the outliers are excluded, then you must be careful in interpreting the results, because the results depend rather delicately on a few outlier data values.

*Remember that if the graph (histogram or frequency polygon) of a distribution is asymmetric, the distribution is said to be **skewed**.* It is also known as **lack of symmetry**. A skewed distribution can be either **positively skewed** (the frequent scores are clustered at the lower end and the tail points towards the higher or more positive scores) or **negatively skewed** (the frequent scores are clustered at the higher end and the tail points towards the lower or more negative scores).

2.5.4: Linearity

Linearity is the assumption that two variables are related in a linear fashion. If variables are linearly related, then when plotted in a scatterplot the data will fall in a straight line or in a cluster that is relatively straight. Sometimes, if the data are not linearly related (i.e., the plot looks curved) the data can be transformed to make the variables linearly related.

2.5.5: Summary of Assumptions

Independence is the most important of these assumptions, and also the most difficult to accommodate when it fails. Constant variance is intermediate, in that non-constant variance can have a substantial effect on our inferences, but non-constant variance can also be accommodated in many situations. Normality is the least important assumption among all, particularly for large sample sizes.

Remember that assumptions can be violated in certain cases during research but not the principles of DOE. This violation of assumptions can be subjective. Our assumptions of independent, normally distributed errors with constant variance are not true for real-world data. However, our procedures may still give us reasonably good inferences, provided that the departures from our assumptions are not too great. Therefore, we assess the nature and degree to which the assumptions are violated and take corrective measures if they are needed.

All the knowledge belongs to God.

CHAPTER 3: DESIGN OF EXPERIMENT

All the knowledge belongs to God.

3.1: DESIGN

There have been four eras in the modern development of statistical experimental design.

3.1.1: Analysis of Variance (ANOVA)

Mendel is known as inventor of design, as he used CRD in his experiments. CRD is classical design which is still being applied in 3rd world countries and has no application in modern world research. However, the agricultural era was led by the pioneering work of **Sir Ronald A. Fisher** in the 1920s and early 1930s. During that time, Fisher was responsible for statistics and data analysis at the Rothamsted Agricultural Experimental Station near London, England. He was hired by Rothamsted Research Station to statistically analyze data and observations collected from continuous research on wheat that was conducted over 70 years to determine the cause of variation in yield.

Making use of extensive long-term field and weather data and notes, Fisher developed ANOVA to estimate and compare the relative contributions of various factors (weather, soil fertility, weed cover etc). After this he published the most famous book: 'Statistical Methods for Research Workers' (1925). Fisher has said that the analysis of variance is merely "a convenient way of arranging the arithmetic". This statement points out that the statistical principles underlying the analysis of variance are quite simple; but the calculations can become quite involved, so that they require careful and systematic arrangement.

Fisher recognized that flaws in the way the experiment that generated the data had been performed often hampered the analysis of data from systems (in this case, agricultural systems). By interacting with scientists and researchers in many fields, he developed the insights that led to the three basic principles of experimental design: randomization, replication, and blocking.

The ANOVA was introduced by Fisher in the context of population genetics (in 1918); he quickly extended the scope (in 1925). His 1918 paper actually introduces the terms "variance" and "analysis of variance". ANOVA can give us an indication that not all the treatment groups have the same mean response, but an ANOVA does not, by itself, tell us which treatments are different or in what ways they differ. ANOVA works by partitioning the total variability in the data into parts that mimic the model. *It quantifies and compares two sources of variation via a particular statistic, an F-ratio for the comparison of two variance estimates.*

3.1.2: Immediacy and Sequentiality

Although applications of statistical design in industrial settings certainly began in the 1930s, the second, or industrial, era was catalyzed by the development of response surface methodology (RSM) by Box and Wilson (in 1951). They recognized and exploited the fact that many industrial experiments are fundamentally different from their agricultural counterparts in two ways: (1) the response variable can usually be observed (nearly) immediately, and (2) the experimenter can quickly learn crucial information from a small group of runs that can be used to plan the next experiment. Box (in 1999) calls these two features of industrial experiments **immediacy** and **sequentiality**. However, the application of statistical design at the plant or manufacturing process level was still not extremely widespread.

It was during this second or industrial era that work on optimal design of experiments began. Kiefer and Wolfowitz proposed a formal approach to selecting a design based on specific objective optimality criteria. Their initial approach was to select a design that would result in the model parameters being estimated with the best possible precision. This approach did not find much application because of the lack of computer tools for its implementation.

All the knowledge belongs to God.

3.1.2.1 OFAT Design

Before 1880, each factor from number of factors in a research was used to be applied by One-way ANOVA i.e. one factor vary and the rest of factors remain constant. This problem led to the development of OFAT (explained in next section). It is one factor at a time (OFAT), sometimes referred to as "**one-at-a-time design** or **one-variable-at-a-time**." It came after ANOVA, when **Karl Pearson** said, it is not possible that only one factor varies every time and the rest of factors remain constant. *OFAT states that change the value of the one factor, then measure the response, repeat the process with another factor.*

Let us observe an example of system research where the effects of k factors on p levels are to be determined. As we mention above, the classical system of experimenting requires each factor to be tested at p levels while others are kept constant at chosen fixed values. The total number of trials to be done by this scheme is:

$$N = k(p-1) + 1$$

The OFAT method consists of selecting a starting point, or baseline set of levels, for each factor, and then successively varying each factor over its range with the other factors held constant at the baseline level. After all tests are performed, a series of graphs are usually constructed showing how the response variable is affected by varying each factor with all other factors held constant. It means find the best sequence first and then go ahead. For example, if we have 7 factors (A, B, C, D, E, F, G) to be evaluated at two levels (1 and 2), then experiment needs to be run 14 times keeping one level constant at a time. For 3 replications, then total 42 runs will be performed. Remember that the response using such methodology cannot be guaranteed as perfect because only single factor varies at a time.

In the experiment of searching optimal temperature and time to maximize broiler yield, this is how the experiment looks using an OFAT method:

- Start with temperature: Find the temperature resulting in the highest yield, between 50 and 120 degrees. Run a total of eight trials. Each trial increases temperature by 10 degrees (i.e., 50, 60, 70 ... all the way to 120 degrees), with time fixed at 20 hours as a controlled variable. Measure yield for each batch.
- Run the second experiment by varying time, to find the optimal value of time (between 4 and 24 hours). Run a total of six trials. Each trial increases time by 4 hours (i.e., 4, 8, 12… up to 24 hours), with temperature fixed at 80 degrees as a controlled variable. Measure yield for each batch.
- After a total of 14 trials, we've identified the max yield happens when temperature is at 80 degrees and time is at 20 hours. As you can already tell, OFAT is a more structured approach compared to trial and error.

There are some advantages of OFAT e.g., the number of the test runs is believed to be close to minimum that can be devised to investigate several factors simultaneously. Similarly, one can readily access the factor effect as the experiment progresses, because only a single factor is being studied at any stage.

However, there may be synergisms (interactions) among the variables that cannot be detected. OFAT assumes that every problem has a single cause. That is why, it does not allow the evaluation of interaction among factors.

All the knowledge belongs to God.

Sometimes, the engineers often conduct experiment in lab in which all factors, but one is strictly under controlled. After the product is moved to production, they often say, 'I don't understand, it worked well in the lab, but it is a disaster in the production, why people in production cannot do as good as engineer in the lab?' The problem is when it is done in lab, engineer create a sterile environment by controlling all but one factor. However, in production, natural variation, noise or production environment is allowed to operate, which create interaction of factors and thus problems (major disadvantage of OFAT, and this is also the reason that most of research papers fail in industrial application). Here comes the factorial experiment to solve this problem i.e. if tea making has 5 factors say sugar, milk, tea, heat and time, then it is not possible that taking one variable at a time and applying ANOVA on it will provide the same result as when all factors will be used altogether. There interaction must give some different result.

3.1.3: Robust Parameter Design

The concepts of Design of Experiments (DoE), alternately known as **robust design** or **variability reduction**. The increasing interest of Western industry in quality improvement that began in the late 1970s ushered in the third era of statistical design. The work of **Genichi Taguchi** had a significant impact on expanding the interest in and use of designed experiments. Taguchi advocated using designed experiments for what he termed robust parameter design, or

- Making processes insensitive to environmental factors or other factors that are difficult to control.
- Making products insensitive to variation transmitted from components.
- Finding levels of the process variables that force the mean to a desired value while simultaneously reducing variability around this value.

Taguchi suggested highly fractionated factorial designs and other orthogonal arrays along with some novel statistical methods to solve these problems.

There were several positive outcomes of the Taguchi controversy. First, designed experiments became more widely used in the discrete parts industries, including automotive and aerospace manufacturing, electronics and semiconductors, and many other industries that had previously made little use of the technique. Second, the fourth era of statistical design began. This era has included a renewed general interest in statistical design by both researchers and practitioners and the development of many new and useful approaches to experimental problems in the industrial world, including alternatives to Taguchi's technical methods that allow his engineering concepts to be carried into practice efficiently and effectively. Third, computer software for construction and evaluation of designs has improved greatly with many new features and capability. Forth, formal education in statistical experimental design is becoming part of many engineering programs in universities, at both undergraduate and graduate levels. The successful integration of good experimental design practice into engineering and science is a key factor in future industrial competitiveness.

3.1.4: Factorial Experiments

In 1990s, it was thought that one-way ANOVA cannot be applied on every design especially with many factors which uses interactions of treatment and whose outputs cannot be separated like inputs after the response has come. This gave rise to the concept of factorial experiment/design, which says that *if two or more factors are involved and they interact to generate the output and are inseparable, then the design will be factorial.* In other words, factorial treatment structure exists when the treatments are the combinations of the levels of two or more combine the factors. In this case one-way ANOVA is not possible.

In summary, all these designs were developed in a time sequence as following:

All the knowledge belongs to God.

$$ANOVA \rightarrow OFAT \rightarrow Factorial\ Design$$

Factorial is the climax of the DOE until new design comes.

3.2: EXPERIMENTAL DESIGN

Design *is an arrangement*. There are several key elements associated with the scientific method, and the concepts and techniques of statistics play a prominent role in all these elements:

1) Observation
2) Hypothesis
3) Design of experiment

Experimental design is the branch of statistics and is also written as design of experiment (DOE). It is a planned approach for determining cause and effect relationships. It literally means a '**systematic plan**'. *Experimental design consists of purposeful changes of inputs (factors) to a process in order to observe the corresponding changes in outputs (response)* (Figure 3.1). It refers to a plan for assigning subjects to experimental conditions and the statistical analysis associated with the plan. This plan is used to collect the data relevant to the problem under the study in such a way as to provide a basis for valid and objective inference about the stated problem. In short, selection of units and assignment of treatments is known as DOE.
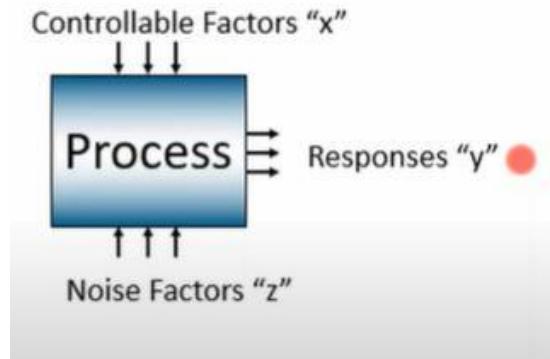
Figure 3.1: Concept of DOE

3.2.1: Characteristics of Design

All scientific research depends on a good design. Not all DOEs are created equal. A good DOE:

- Should produce optimum output with minimum expenditure.
- Should have randomization and be replicable.
- Give max information.
- Be easy to implement.
- Should provide reliable findings: should provide logical and critical outcomes.
- Should have the capacity to evolve with the passage of time (a design is not suitable for all the times e.g., rise and fall of Amazon)
- Should avoid systematic errors.
- Should be precise and ignore false information/effects.

Why do most of postgraduate research papers have no practical application in industry? Because there exist problems in DOE.

All the knowledge belongs to God.

3.2.2: Components of DOE

The DOE mainly has three components:

- Planning of experiment
- Obtaining the relevant information
- Statistical analysis of information and drawing the inference.

3.2.3: Applications of DOE

Experimental design is a critically important tool in the scientific and engineering world for improving the product realization process. Critical components of these activities are in new manufacturing process design and development, and process management. The application of experimental design techniques early in process development can result in:

- Improved process yields
- Reduced variability and closer conformance to nominal or target requirements
- Reduced development time
- Reduced overall costs.

3.2.4: Types of Designs

There are different types of designs i.e., classical and modern designs. Broadly, there are two types of designs:

- Systematic
- Random

ANOVA techniques are suitable for only random designs. There are three types of random designs:

1) CRD: completely randomized design. It is known as one-way ANOVA in SPSS as only factor is tested in it.
2) RCBD: randomized complete block design. It is also known as two-way ANOVA in SPSS as two factors are tested in it.
3) LSD: Latin square design or three-way ANOVA in SPSS as three factors (or one factor and two blocking) are involved.

There is a difference between ANOVA and design. Design is a plan but there is no execution of the plan involved in it. However, ANOVA contains design + counterpart (execution, outcome, analysis, significance establishment). In other words, ANOVA contains all steps from hypothesis to outcome including planning.

3.3: PRINCIPLES OF DOE

Any research which is non-deterministic, is based on three experimental design principles:

1) Randomization
2) Replication
3) Local control

Some authors write 'blocking' as 3rd principle. These 3 principles should be followed during the DOE at each and every level.

3.3.1: Randomization

This is the first principle of DOE. *It is unbiased random allocation of the treatments among the EUs.* The assignment of treatments to EUs is done by using a chance mechanism and implies that every possible allotment of treatments has the same probability. The purpose of this is to remove the bias and other source of extraneous variation, which are not controllable. Randomization is opposite to 'convenient sampling' and ensures the true representation of

All the knowledge belongs to God.

population. If not followed, the response will differ and be non-represented. Random assignment of treatments has reasons for useful consequences:

- It protects against confounding.
- It can form the basis for inference
- It helps in valid estimation of experimental error
- Randomization also makes the experiment free from systematic errors
- Independence of errors in normality assumption is also achieved

Randomization can be either physical or numerical. **Physical randomization** is achieved via an actual physical act that is believed to produce random results with known properties, i.e. coin tosses, card draws from shuffled decks, physical rolls of a die, and tickets in a hat. **Numerical randomization** uses numbers taken from a table of "random" numbers or generated by a "random" number generator in computer software. However, the numbers produced by the software package are from an algorithm; if you know the algorithm you can predict the numbers perfectly. That is why, they are technically **pseudorandom numbers**. That is why always randomly allocate the levels of factors and replicate them accordingly.

3.3.2: Replication

To have valid and accurate estimation of means and variances due to different sources of variations, one needs to apply the treatments into more than one EU. *Replication is the assignment of individual treatments to multiple EUs.* In other words, it is the **repetition of treatment.** An individual repetition is known as **replicate**. Remember that either toss 1 coin for 3 times or toss 3 coins for 1 times, it will be the same for statistics.

Replication reduces the chance of unknown error in the treatment and increases the confidence. A treatment is repeated to have a more reliable estimate than what would be possible from a single observation. Because there exist some variation in EUs, and these units cannot be physically identical. Thus replication contributes to statistical analysis of the experimental data in two ways:

- It provides the data from which to estimate the inherent variability of EUs.
- It influences the precision with which the treatments can be compared.

The variability of responses over these replications, which has the statistical term, experimental error, provides a yardstick against which we measure differences among treatments. In other words, replication is used to secure more accurate estimate of the experimental error. Replication should be enough according to the budget of the researcher. Remember to ensure replication on three levels:

1) Stage 1 replication: replication of factors
2) Stage 2 replication: replication of levels of factors
3) Stage 3 replication: repeat the whole process.

3.3.3: Local Control or Blocking

All extraneous sources of variations cannot be removed by randomization and replication. Researchers need to choose a design in such a way that all extraneous sources of variation are brought under control. That is why we use local control, which is the amount of balancing, blocking and grouping of the EUs. The term local control should not be confused with 'control treatment/group'.

Local situations for different experiments vary, and one needs to take care of such variation during experimentation so as to minimize the error. *Local control, simply, is the technique which helps in the reduction of experimental error, providing due consideration*

All the knowledge belongs to God.

to the information available under the local conditions where the actual experiment is conducted. Using the locally available information like shape of the experimental plot, direction of the field, its fertility gradient, slope, and other conditions nearby the plot designs are to be framed in such a way so as to reduce the experimental error.

Local control is like constraints of your experiment. Experimental protocols need to be established and implemented that prevent freelancing (and much more subtle modifications) and protect the integrity of the design and subsequent data analysis. These are the constraints which are out of the control of researcher and must be explicitly mentioned in the research methodology and response. For example, if a researcher has 3 incubators of different origin and age, then it may be possible that they might not provide same set of conditions in the recommended time frame. This should be mentioned as local control in the research, because researcher cannot purchase or arrange 3 incubators with the same set of performance indicators.

3.3.3.1 Blocking

Some authors mention blocking as third principle of DOE instead of local control. However, it is almost the same as local control. Blocking is used in the research where instruments are involved i.e. if we have 3 incubators for egg hatching, we will have to ensure 3 blocks at least. Normally it is used when we move from CRD to RCBD.

Blocking is the assignment of treatments within multiple groups of EUs. Conventional terminology is to refer to groups of EUs as "blocks." So, a block is a group of EUs homogeneous in nature. In experiments conducted with animals, blocking may be done according to the age, weight, sex, etc. group of animals, where animals of the same age or similar weight or same sex may form the blocks.

Blocking is a design technique used to improve the precision with which comparisons among the factors of interest are made. Often blocking is used to reduce or eliminate the variability transmitted from nuisance factors—that is, factors that may influence the experimental response but in which we are not directly interested. For example, an experiment in a chemical process may require two batches of raw material to make all the required runs. However, there could be differences between the batches due to supplier-to-supplier variability, and if we are not specifically interested in this effect, we would think of the batches of raw material as a nuisance factor.

There are many ways that the EUs selected for an experiment can be organized or structured. The simplest situation is to have one group of homogeneous EUs (meaning similar, not identical), such as individual plots of land in a field or a garden in which the soil quality is presumed or known to be essentially the same throughout. Alternatively, an experiment can have multiple groups of EUs. The groups might differ in some recognizable way, but within each group, the EUs would be relatively homogeneous. An example is plots of land in different fields or regions.

Blocking also defines the scope of the conclusions that can be drawn from an experiment. In agriculture, a developer of a new variety of corn wants to know whether that variety outproduces other varieties in a (reasonably) wide variety of soils and growing conditions. Thus, the experiment will be blocked by location, with the locations selected to span the desired range of conditions. A new variety that excels only in a very limited set of conditions is not as marketable as one that excels in a broad set of conditions.

3.4: DESIGN FORMULATION

Any design formulation contains the following things to be elaborated:
1) Hypothesis

All the knowledge belongs to God.

2) Model
3) Layout
4) Name of the design

3.4.1: Hypothesis

A hypothesis is simply a statement about one or more populations. *A **statistical hypothesis** is an assertion/statement about the probability distribution of population characteristic(s) which is (are) to be verified based on sample information.* For example, the statement about the students of particular university is that the IQ of the students is 0.9 in 1.0 point scale or the average milk yield of a particular breed of cow is 3500 / for liter per lactation. Now on the basis of sample observations, we are to verify the statements that the IQ of the students 0.9 in 1.0 point scale or not, and average milk yield of a particular breed of cow is 3500 / for liter per lactation or not. In short, hypothesis is an **unproven statement**.

There are normally **two types of hypotheses** to be formulated for test: null hypothesis (H_o) and alternative hypothesis (H_1).

H_o is called 'null hypothesis and is already established. However, H_1 is known as 'alternative hypothesis' and is always a new one (to be established). Technically, *unbiased/unmotivated statistical hypothesis whose validity is to be verified for possible acceptance or rejection based on sample observations is called **null hypothesis**.*

*The statistical hypothesis which differs from the null hypothesis is called the **alternative hypothesis**.* In the above two examples, the statements that (a) the students of a particular university has the IQ 0.9 in 1.0 point scale and (b) the average milk yield of the particular breed of cow is 3500 / of liter per lactation are null hypotheses, whereas any hypothesis like (c) IQ of the students is not 0.9 in 1.0 point scale or (d) the average milk yield of the particular breed of cow is not 3500 / of liter per lactation or the average milk yield of the particular breed of cow is less than 3500 / of liter per lactation or the average milk yield of the particular breed of cow is more 3500 / of liter per lactation etc. are the examples of alternative hypothesis.

Write only one hypothesis (H_1) on which you are working on, in thesis to save time and space. As, **researcher hypothesis is always H_1.** If there is only one factor involved, then only one hypothesis (H_1) will be written in paper. In case of two factors, write both hypotheses along with interaction separately and all should be H_1.

Axiom is alternative to hypothesis and is **taken as true without any proof**; however, axiom is the same as hypothesis but becomes a part of belief. It is not written in research.

3.4.2: Model

Model means model of response data. *Model for data is a specification of the statistical distribution for the data.* For example, the number of heads in 10 tosses of a fair coin would have a binomial (10, 0.5) distribution, where 0.5 gives the probability of a success and 10 is the number of trials. In this instance, the distribution depends on two numbers (10 and 0.5), called **parameters**: the number of trials and the success probability respectively. For 10 tosses of a fair coin, we know both parameters.

The objectives of the experiment can often be described as deciding which model is the best for description of the data and making inferences about the parameters in the models. Statistical model of an experiment has 3 components:

$$Response = f\ (explanatory\ variable) + Error$$

All the knowledge belongs to God.

It can also be written as:
$$Response = Predicate\ Structure + Error$$

Response is a dependent variable. It is better to have this variable either as ratio or interval but mostly ratio is recommended. Explanatory variable/predicator variable (refers to treatments and blocking) should belong to ordinal factor. However, response always falls under quantitative factor category and predicator structure is always qualitative factor. Similarly, predicate structure is independent variable and is what all variables of a design/model, it comes in right side of model. This right side is controllable variable (predicate structure).

Remember that even one variable cannot be called as predicate alone, it will be called as predicate structure because even one predicator can make uncontrollable structure by uncontrollable combinations. With an example of effect of temperature on broiler weight gain, model can be written as:
$$Weight\ gain = Temperature + error$$

In case of CRD, only one factor is written in predicate structure; however, in case of two factors, these are written with plus sign (+) to mark the addition, i.e., in RCBD when writing factor and blocking i.e., effect of temperature on 2 strains of broilers (Ross and Cobb):
$$Weight\ gain = (Temperature + strains) + error$$

If there is interaction involve in them, predicate structure is written two times, first with plus sign (+) to show the individual main effects and then with multiplication sign ($\times$) to show the interaction effect i.e. effect of temperature with age on broilers:
$$Weight\ gain = (Temperature + age) + (temperature \times age) + error$$

Remember that error is not mentioned specifically. Just write the word error. Even error is not discussed in the research paper or thesis.

3.4.3: Layout

Systematic arrangement of EUs in the experimental field and allocation of treatments in EUs according to the requirement of the specific design is called layout. It is mainly concerned in field experimentation. Layout varies from design to design and type of experiments.

For example, if the whole experimental field is homogeneous in nature, then it is divided into as many numbers of EUs as the total number of replications required to allot all the treatments with respective number of replications. On the other hand, if the EUs are heterogeneous in nature, then at the first step of layout, blocks are required to be framed, and then treatments are to be allocated (Table 3.1). Layout is usually in the tabular form of a DOE showing the replication and treatments to be tested over EUs.

Table 3.1: Layout of DOE

Factor (treatments)	Treatment 1	Treatment 1	Treatment 1
Replicates (EUs)	A	B	C
Subject per experimental units	10	10	10

3.4.4: Name of Design

It is the design to be implemented on the hypothesis i.e., CRD, RCBD or LSD. The reason is not mentioned until asked. Similarly, we cannot write which ANOVA is to be applied because

All the knowledge belongs to God.

no execution is involved at this stage, especially when writing synopsis. However, when writing thesis, ANOVA type can be written because now the execution has been done. Though in synopsis, ANOVA type can be written in future form.

3.4.4.1 No-Design Example

Remember that if response is nominal (qualitative), there is no design applicable. Go through some of the following examples:

A university want to know the satisfaction of postgraduate students (5 students from each of 5 degrees) during their course work. She designed an online Performa for evaluation and asked to answer the questions relative to their class, teacher and subject as following:

- Are you satisfied with the class environment?
- Was the teacher punctual?
- Did this subject increase your knowledge?

After the evaluation students answer was as noted (Table 3.2). Recognize the design and do appropriate analysis.

Table 3.2: Teacher evaluation data of students.

Questions	Answers	Poultry	Dairy	Zoology	Livestock	Genetics
Satisfaction?	Yes	4	2	3	2	3
	No	1	2	2	0	2
	Not answered	0	1	0	3	0
Improvement?	Yes	4	3	2	2	5
	No	0	0	3	3	0
	Not answered	1	2	0	0	0
Class environment?	Good	2	0	1	1	1
	Bad	3	1	4	4	3
	Not answered	0	4	0	3	1
Knowledge?	Increased	3	2	2	3	1
	No improvement	2	2	3	2	2
	Not answered	0	1	0	0	2
Teacher cooperation?	Yes	3	3	3	1	1
	No	1	2	2	2	4
	Not answered	1	0	0	2	0
Course content	Updated	2	2	1	3	0
	old	3	2	2	2	5
	Not answered	0	1	2	0	1

There are following variables in it:

- $Response = Nominal\ (yes; no; not\ asnwered)$
- $Factor = Question\ (6\ types\ as\ 6\ levels)$
- $Experimental\ Units = Degrees$
- $Subjects = Students$

So, no design is applicable

All the knowledge belongs to God.

Similarly, Table 3.3 comes from a hypothetical study of factors influencing the primary feed choice of desi chickens. The study captured 219 chickens in four areas of Pakistan. The nominal-scale response variable is the primary feed type, in volume, found in chicken's stomach. Table 3.3 classifies the primary feed choice according to the area of capture and the weight of the chicken. Here, weight is binary, distinguishing between non producer (weight $\leq$ 2.3 kg) and producer chickens.

Table 3.3: Primary feed choice of chickens.

Area	Weight (kg)	Primary Feed Choice				
		LSB	Desi	Misri	Golden	Mixed
Islamabad	Less than 2.3	23	4	2	2	8
	Greater then 2.3	7	0	1	3	5
Rawalpindi	Less than 2.3	5	11	1	3	5
	Greater then 2.3	13	8	6	1	0
Sarghoda	Less than 2.3	5	11	2	1	5
	Greater then 2.3	8	7	6	3	5
Gilgit	Less than 2.3	16	19	1	2	3
	Greater then 2.3	17	1	0	1	3

The variables are:
- $Response = choice\ of\ feed\ (nominal)$
- $Factor = areas\ (4\ types\ as\ 4\ levels)$
- $Factor\ 2 = weight\ of\ chicken\ (2\ levels: < 2.3\ kg\ and > 2.3\ kg)$

None of the design is possible here.

All the knowledge belongs to God.

CHAPTER 4: COMPLETELY RANDOMIZED DESIGN

4.1: INTRODUCTION

When there is only one factor involved in it without any constraint, or the whole experimental area or all the EUs are homogenous in nature, then one can think of completely randomized design (CRD). In this design, out of the 3 basic principles of the DOE, only 2 principles, viz., replication and randomization, have been used. The third principle, (local control), is not used as such the minimization of error is not there.

4.1.1: CRD Randomization

We may randomly allocate subjects to treatments as follows (Figure 4.1). Suppose we have 16 subjects available to participate in an experiment in which we wish to compare four drugs. Subjects can be numbered from 1 through 16. Go to a table of random numbers and select 16 consecutive, unduplicated numbers between 1 and 16.

Now randomly allocate them using random number generation e.g.,16, 09, 06, 15, 14, 11, 02, 04, 10, 07, 05, 13, 03, 12, 01, and 08. Allocate subjects 16, 09, 06, and 15 to drug A; subjects 14, 11, 02, and 04 to drug B; subjects 10, 07, 05, and 13 to drug C; and subjects 03, 12, 01, and 08 to drug D. It is emphasized that the number of subjects in each treatment group does not have to be the same.
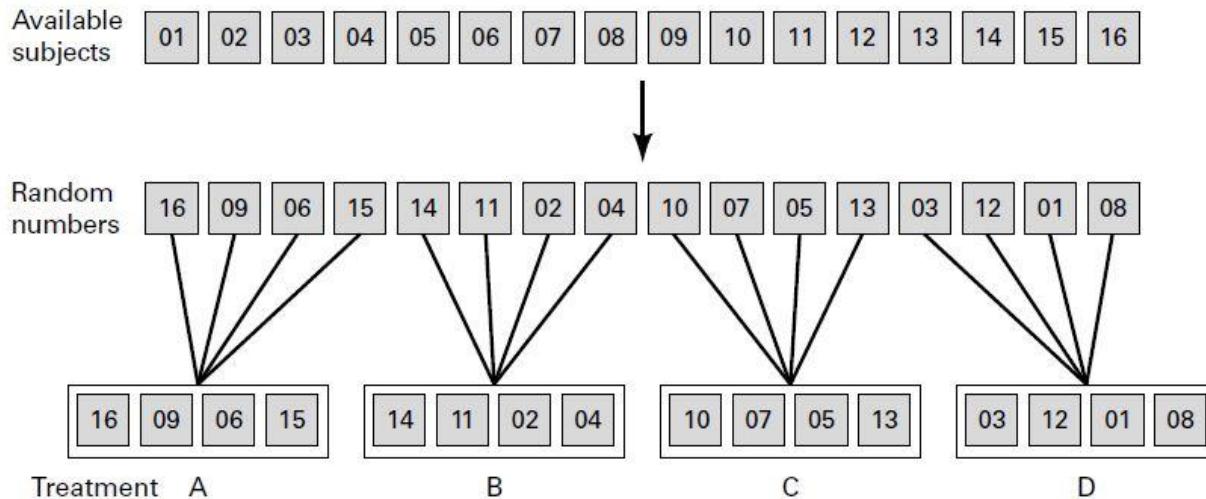
Figure 4.1: CRD randomization procedure.

4.1.2: Merits And Demerits Of CRD

CRD is the simplest among all DOEs. It is the only basic design where one can have flexibility of adopting different numbers of replications for different treatments. When the experimenter comes across with the problem of varied availability of experimental materials, the flexibility of different replications for different treatments becomes very useful. Missing data does not provide potential threat so long there are a few observations corresponding to each treatment. This is possible only because of the design flexibility to handle different replications for different treatments. Compared to other basic designs, CRD can be used in irregular shaped experimental plot.

However, though CRD is most suitable for laboratory or greenhouse experimental condition because of homogenous EUs, it is very difficult to have homogenous EUs in large number under field condition. In CRD only two basic principles of the DOE are used. The principle of "local control" is not used in this design which is very efficient in reducing the experimental error. As experimental error is more compared to other basic DOEs.

All the knowledge belongs to God.

4.2: FORMULATION OF CRD

Suppose we have group of 50 animals (homogenous) and we want to test one factor (feed choice) with 5 levels of factor (feed 1, 2, 3, 4 and 5) on the weight gain (yield) of these animals. The plan will be as following:

We'll make 5 feed levels or treatments (homogenous), because we have 5 choices of feed to be tested (first level of replication), to which will be assigned 50 animals randomly: 10 animals per replication (second level of replication). One choice of feed will be given to each replicate from five feeds. This process will be repeated as a whole as required (third level of replication). If the process is repeated, any animals form level one in factors (choice of feed) may have chance to be assigned to any choice of feed and there is no compulsion unlike RCBD where the whole block (10 animals) should be assigned altogether to any choice of feed. In other words, for CRD, if feed 1 has block of 10 animals in first process, then in next process of replication of the process, it might get either the same 10 animals or any combination of animals based on draw (complete randomization).

Remember that we cannot have 2 or more factors in CRD as some of authors have argued by naming them as factorial CRD. If two factors are involved, either it will be RCBD or factorial (if interaction is involved). Factorial CRD is possible only when two-way ANOVA is converted into one-way ANOVA by de-coding factor (a term explained in later chapters).

4.2.1: Hypothesis

Let's take first scenario. Hypotheses from above mentioned example can be written as:

$$H_o = all\ feeds\ are\ equal\ or\ have\ same\ response.$$
$$H_1 = At\ least\ one\ feed\ is\ not\ equal\ to\ others\ in\ response.$$

The H_1 can also be written in another way if manufacturer is experimenting himself, say its feed is Feed 1. Then the H_1 will be:

$$H_1 = Feed\ 1\ is\ not\ equal\ to\ others\ in\ response.$$

It is highly discouraged to say (for H_1), 'all feeds are non-equal', because there might be possible that any two feeds give the same outcome. In this way, our alternative hypothesis will be a wrong statement.

Now take another scenario and suppose we have already one best feed in the market and want to test new feed in its comparison along with other three feeds, then the hypotheses will be:

$$H_o = Feed\ 3\ has\ the\ best\ response$$
$$H_1 = Our\ feed\ (feed\ 1)\ is\ better\ than\ feed\ 3$$

4.2.2: Model

$$Weight\ gain\ (yield) = Feed\ choice + error$$

4.2.3: Layout

It will be: $5^1 CRD = (5\ treatments)$.

All the knowledge belongs to God.

Table 4.1: Layout of CRD.

Feed choice (factors)	Feed 1	Feed 2	Feed 3	Feed 4	Feed 5
Replicates	A	B	C	D	E
Broilers per replicate	10	10	10	10	10

4.2.4: Name of Design
CRD will be applied.

4.3: DESIGN RECOGNITION

4.3.1: Example 1
The data on anxiety obtained on students of DVM, Poultry, and Zoology students is shown (Table 4.2). Score was recorded from 0-100. The higher the score is higher the anxiety level. Find in which degree anxiety is higher. Discuss the findings at 5% level.

Table 4.2: Anxiety of students.

SN	DVM	Poultry	Zoology
1	22	25	25
2	21	20	30
3	21	19	28
4	23	20	20
5	22	16	26
6	23	18	29
7	21	21	33
8	24	16	33
9	22	17	27

Variables included are:
- $Dependent\ Variable = anxiety$
- $Factor = Degree$

The hypothesis need to be tested will be:
$$H_1 = Change\ in\ degree\ has\ an\ effect\ on\ anxiety$$
Layout has already mentioned in the example. It is a CR design.

4.3.2: Example 2
A university is studying the horizontal jump (feet) of male students in a playing ground and obtained five students in four different degrees: DVM, Poultry, Zoology and BBA (Table 4.3). She wants to test whether the physical strength of the students is the same in all three degrees.

All the knowledge belongs to God.

Table 4.3: Horizontal jump of students in feet.

DVM	7	8	6	11	8
Poultry	11	12	12	9	8
BBA	7	7	8	8	8
Zoology	6	6	7	7	8

It is CRD and variables are:
- $Factor = Degree$
- $Levels\ of\ factor = 3\ (DVM;\ Poultry;\ BBA; Zoology)$
- $Response = Horizontal\ jump\ (feet)$

The hypothesis will be:

$H_1 = At\ least\ one\ of\ the\ degree\ has\ different\ jump\ length\ from\ others$

4.3.3: Example 3

Researchers were interested in the acid treatment effect on growth rate of broilers. They created three water treatment groups in an experiment: low acid, high acid and control. The response variable in their experiment was the average weight of the broilers in a poultry house after 35 days of growth. The observational unit was a pen, rather than individual bird. The data are given below (Table 4.4), do appropriate analysis.

Table 4.4: Broiler growth under three water treatments.

Water treatments		
Low acid	**High acid**	**Control**
1.58	1.1	2.47
1.15	1.05	2.15
1.27	0.5	1.46
1.25	1	2.36
1	1.5	1

It is CRD and variables included are:
- $Factor = acid\ treatment\ (3\ levels)$
- $Response = weight\ of\ broilers$

The hypothesis needs to be tested is:

$H_1 = At\ least\ one\ of\ the\ treatment\ gives\ different\ broiler\ weight$

The layout is already mentioned in the example.

All the knowledge belongs to God.

CHAPTER 5: RANDOMIZED COMPLETE BLOCK DESIGN

5.1: INTRODUCTION

A subtype of **randomized block design (RBD)**, it is the simplest of all field designs which uses all the basic principles of DOE. It was developed about 1925 by **R. A. Fisher**, who was seeking methods of improving agricultural field experiments. When there are two factors involved or one factor with one constraint (blocking), then it will be a RCBD. Remember that every RCBD is 2-way ANOVA, but every 2-way ANOVA is not RCBD. We will call the design as RCBD when both factors will have same levels, otherwise it will be called as 2-way ANOVA only (which is also known as **factorial ANOVA**).

For example, a university wished to compare three methods of teaching the students to check the improvement in grades. She felt that the rate of learning would be different for students of different degrees and wished to design an experiment in which the influence of degree could be considered. Three students in each of five degrees were selected to participate in the experiment, and one student in each degree was randomly assigned to each of the teaching methods (Table 5.1). The methods of instruction constitute our three treatments, and the five degrees are the blocks.

Table 5.1: Grades of students.

Degrees	Teaching Method		
	Modern	Classical	Mixed
Poultry	7	9	10
DVM	8	9	10
Zoology	9	9	12
BBA	10	9	12
Botany	11	12	14

Blocking is done across the factor gradient (may be horizontal or vertical), and each block should contain as many EUs as the number of treatments in the experiment. *As each and every block contains all the treatments, the blocks are known to be complete, and as such the design is also known as randomized 'complete block' design (RCBD).* However, blocking designs are not completely randomized designs. Because, these use some kind of **'restricted randomization'**.

RCBD opened ways for new designs and removed a myth that only one design (CRD) existed. It is only used when we have heterogeneous EUs (unlike CRD). However, if there are two factors, it will be factorial RCBD.

5.1.1: Difference between RCBD and CRD

The main difference between these two is the material (EUs) homogeneity and number of factors. If the material is homogeneous then CRD will be applied, otherwise for heterogeneous material RCBD is used. Because error cannot be controlled in CRD, to avoid this experimental error we use blocking in RCBD (blocking reduces experimental error) and then place homogeneous material in it. Afterwards, all the blocks are given whole set of treatments. For example, if we want to test a feed effect on different temperatures on broiler growth, there will exist certain variations in environment. To control this variation, we will split them into blocks (according to number of temperatures) and apply all treatments (feed) on EUs within each block. Similarly, if there is only one factor to be used,

All the knowledge belongs to God.

then it will be CRD (without blocking), if there is one factor with constraint (blocking) or simply two factors, then it must be RCBD.

Furthermore, it is randomization of 'complete block' unlike the CRD. Because during third level of replication, there will be the randomization of a 'complete block' and not the EUs alone. There is no chance that the any EU from one level of factor (say feed 1) could be assigned to other level of factor (say feed 2). If so, it will be all EUs together within each block at a time. We cannot fix whole block of EUs in CRD. Because, in CRD, there may or may not be new set of EUs in one level of factor. In other words, composition of the block remains the same in RCBD and may not be same in CRD. This is also the reason that replication is always insignificant in RCBD. As any treatment will produce the same result with block A as any other block. For example, if Feed 1 has best result on block 'A' first time, then upon repetition of the whole experiment (third level of replication), it should have the same response on any block upon reassignment.

5.1.2: Merits And Demerits of RCBD

It is the simplest of all block design and flexible: any number of replication and treatments can be run. It uses all the 3 principles of DOE and takes care of heterogeneity of EUs and control the extraneous source of variation. The layout is very simple, experiment can be set easily and is more efficient compared to CRD. RCBD can be used in case of 'missing observation' among replicates.

However, RCBD is a complete block design; each and every block contains all the treatments. If the number of treatments increases to a great extent, block size also increases simultaneously. It becomes very difficult to have a greater homogeneous block to accommodate more number of treatments. In practice, the number of treatments in RCBD should not exceed 12. Each and every treatment is repeated equal number of times in RCBD. As such like CRD, the flexibility of using different replications for different treatments is not possible in RCBD. The missing observation, if any, is to be estimated first and then analysis of data to be taken. RCBD takes care of heterogeneity of experimental area in only one direction.

5.2: FORMUALTION OF RCBD (1 FACTOR AND 1 CONSTRAINT)

Suppose we want to check the effect of different feeds (A, B and C) on two broiler strains (Ross and Cobb) for weight gain (yield). There are 30 broilers in each strain.

We have one factor (choice of feed) and one constraint (strains of broiler). We will assign 10 broilers per EU in every block (3 replicates per EUs per block). So, it will be RCBD (or two-way ANOVA) and not the factorial. If we had taken two factors (say feed choices 'with' temperatures) in different strains, then it would have been factorial RCBD.

5.2.1: Hypothesis

If there is one factor with one constraint, then there will be two hypotheses as no interaction is involved. When there are two factors, then there would be three hypotheses as one for each factor (main effects) and one for interaction effect. In the above stated problem, the hypothesis will be as following:

$$H_o = Both\ feed\ choices\ have\ the\ same\ response$$
$$H_1 = At\ least\ one\ feed\ is\ not\ equal\ to\ others$$

$$H_o' = Both\ strains\ have\ equal\ response$$
$$H_1' = Change\ in\ strains\ has\ an\ effect\ on\ weight\ gain$$

We will not write the third hypothesis as there is no interaction involved. It is written only in case of factorial experiments (factorial RCBD).

All the knowledge belongs to God.

5.2.2: Model

Predicate structure will contain both factor and blocking:

$$Weight\ gain\ (yield) = (Feed\ + strains) + Error$$

The plus sign shows that there is no interaction involved in it.

5.2.3: Layout

It will be $(2 \times 3\ RCBD = 6\ replicates)$ under two strains as following:

Table 5.2: Layout of RCBD (1 factor and 1 constraint).

		Strains	
		Cobb	**Ross**
	A	$A + Cobb$	$A + Ross$
Feed choice	**B**	$B + Cobb$	$B + Ross$
	C	$C + Cobb$	$C + Ross$

Every EU consists of 10 broilers. Note that all EUs in each block is given all treatments (feed choices) which is a complete randomization within a block. In the simplest situation, the number of treatments equals the number of EUs in a block (3 EUs and 3 feed choices), so each treatment is assigned to only one EU in each block. Remember that the RCBD can have unequal replication (of treatment assignments within blocks).

5.2.4: Design Name

It is RCB design.

If there is interaction involved, we can write, RCBD under factorial or 'factorial RCBD', but better is to write factorial design alone. Remember that you cannot determine the design used in an experiment just by looking at a table of results, you must know the randomization.

5.2.5: Establishing the Significance

In the event of significance test(s), one rejects the null hypothesis (hypotheses), resulting in the decision that there exist significant differences among the treatments and among the replications. Thus, there is a need to find out which pairs of the treatments are significantly different from each other, and which treatment is the best treatment or what is the order of the treatments with respect to the particular character under consideration? Similarly, there is a need to find out which pair of replications differs significantly among themselves.

5.3: FORMUALTION OF RCBD (TWO FACTORS & NO CONSTRAINT)

Suppose we have 100 broilers and want to check the effect of two proteins (A and B) and two energy levels (High and Low) on weight gain (yield).

We will not use blocking in this case because we have two factors (choice of protein and energy). That is why, it will be RCBD under factorial (or two-way ANOVA for the purpose of analysis).

All the knowledge belongs to God.

5.3.1: Hypothesis

Here there are two factors, then there will be at least three hypotheses: one for each factor (so-called main effect) and one for interaction (so-called interaction effect) as following:

$$H_o = Both\ the\ protein\ levels\ have\ the\ same\ response$$
$$H_1 = Change\ in\ protein\ has\ an\ effect\ on\ weight\ gain$$

$$H'_o = Both\ energy\ levels\ have\ equal\ response$$
$$H'_1 = Change\ in\ energy\ has\ an\ effect\ on\ weight\ gain$$

$$H''_o = All\ the\ interactions\ have\ equal\ response$$
$$H'_1 = There\ exist\ an\ interaction\ between\ protein\ and\ energy$$

5.3.2: Model

Predicate structure will contain both factors and interaction:
$$Weight\ gain\ = (Protein + energy) + (protein \times energy) + error$$

Notice the repetition of factors in predicate structure and the use of multiplication sign in second predicate which indicate the presence of interaction.

5.3.3: Layout

It will be ($2^2 factorial\ RCBD = 4\ treatments$). Twenty-five broilers will be allocated per treatment ($25 \times 4\ = 100\ broilers$) as following:

Table 5.3: Layout of factorial RCBD.

		Proteins	
		A	**B**
Energy	**High**	$High + A$	$High + B$
	Low	$Low + A$	$Low + B$

5.3.4: Design Name

It is factorial RCBD or two-way ANOVA (**factorial ANOVA**). Remember that you cannot determine the design used in an experiment just by looking at a table of results, you must know the randomization.

5.4: DESIGN RECOGNITION

5.4.1: Example 1

Recognize the design from the following Figure 5.1.

All the knowledge belongs to God.
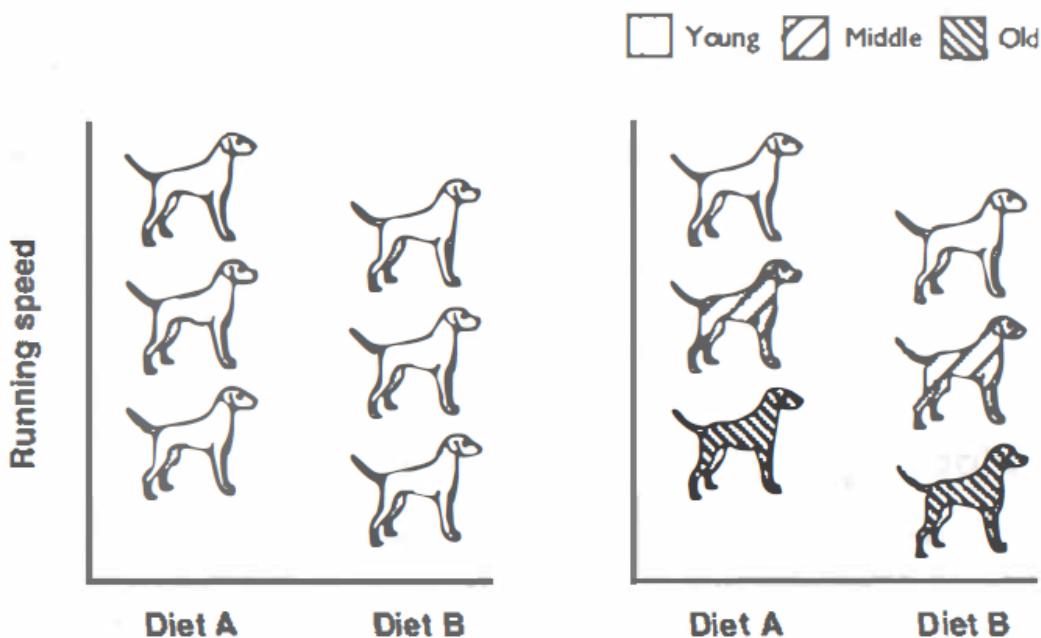
Figure 5.1: Effect of diet on dog's running speed.

Variables included in first figure (Figure 5.1) are:
- $Factor = diet$
- $Response = running\ speed$
- $EUs = dogs$

Similarly, variables in second figure (Figure 5.1) are:
- $Factor = diet\ (only\ one\ factor)$
- $Constraint = varieties\ of\ dog\ (3\ types\ or\ 3\ levels)$
- $EUs = dogs$

The hypothesis needs to be tested in first figure is:

$H_1 = At\ least\ one\ of\ the\ diet\ produces\ different\ running\ speed\ from\ others$

The hypotheses need to be tested in second figure are:

$H_1' = At\ least\ one\ of\ the\ diet\ produces\ different\ running\ speed\ from\ others$

$H_1'' = At\ least\ one\ of\ the\ variety\ produces\ different\ running\ speed\ from\ others$

The first figure is a CR design, and the second figure is RCBD (two-way ANOVA) but no factorial, because three varieties of dog has been taken a constraint/blocking and not the factor.

5.4.2: Example 2

Before new poultry feed is marketed, it is common practice to test them first in research birds e.g., chickens or other respective birds for which the feed has been formulated. In part of one study, a new feed under investigation was given to 4 male and 4 female chickens, at doses 50g/kg and 100g/kg for 7 days. Growth (grams) was measured from samples in order to analyze the effect. The design of this experiment allows for the

All the knowledge belongs to God.

investigation of the interaction of two factors: sex of the chicken and feed amount. Data is shown here (Table 5.4):

Table 5.4: Growth effect under feed and gender.

Feed amount	Male	Female
50	191	150
	154	127
	194	152
	183	105
100	80	141
	49	153
	78	171
	71	197

Variables included here are
- $Factor1 = drug\ dose\ (2\ levels)$
- $Factor\ 2 = gender\ (male\ and\ female)$
- $Response = alkaline\ phosphate\ level$

The hypotheses will be:

$H_1 = Change\ in\ drug\ dose\ has\ an\ effect\ on\ alkaline\ phosphate\ level$

$H_1' = Change\ in\ gender\ has\ an\ effect\ on\ alkaline\ phosphate\ level$

$H_1'' = There\ exist\ an\ interaction\ between\ gender\ and\ drug\ dose$

It is a factorial RCBD or factorial ANOVA.

5.4.3: Example 3

Recognize the design from following table:

Table 5.5: Problem layout for identification.

		Contamination levels					
		1	2	3	4	5	6
Dilution	3	1.87506	1.74036	1.79934	2.02119	1.79934	1.59106
	6	1.38021	1.36173	1.25527	1.39794	1.20412	1.25527
	9	0.60206	0.90309	0.95424	1.00000	0.60206	060206

The variables included here are:
- $Factor\ 1 = dilution\ (3\ levels)$
- $Factor\ 2 = contamination\ (5\ levels)$
- $Response = written\ inside\ the\ table\ boxes$

The hypotheses will be:

$H_1 = Change\ in\ contamination\ level\ has\ an\ effect$

$H_1' = Change\ in\ dilution\ level\ has\ an\ effect$

All the knowledge belongs to God.

$$H_1'' = There\ exist\ an\ interaction\ between\ dilution\ and\ contamination$$

It is a factorial RCBD or factorial ANOVA because two factors have been given in above table.

CHAPTER 6: LATIN SQUARE DESIGN

6.1: INTRODUCTION

RCBD allows us to block on a single source of variation in the responses. There are experimental situations with more than one source of extraneous variation, and we need designs for these situations. Such designs are referred to as designs for two-way elimination of heterogeneity, or **row-column designs.** Latin Square Design (LSD) is type of one such designs.

When there are three factors involved or one factor with two constraints, then it will be LSD. If 3 factors have same levels then it will be named LSD, otherwise three-way ANOVA. Remember that every LSD is three-way ANOVA, but the inverse is not true e.g., 3^3 is LSD but $3 \times 4 \times 2$ design should be named three-way ANOVA.

6.1.1: LSD Layout

A simple example is the assessment of the feed performance of broiler types. Different types of feeders can be fitted in poultry house e.g., four feeder for each of broiler type (Table 6.1). There may be differences in performance between the four feeds (1, 2, 3 and 4) which are consistent for all strains of broilers (I, II, III and IV). There will certainly be overall performance differences between broilers. To compare four brands of feeders (A, B, C, D) using four broilers, we would like to allocate feeder for each broiler strain so that each feeder is tested on each feed and also for each strain of broiler. Can this be achieved?

Table 6.1: Typical LSD layout: broilers; feed and feeders.

		Broiler strains			
		I	II	III	IV
	1	A	B	C	D
Feed	2	B	D	A	C
	3	C	A	D	B
	4	D	C	B	A

Construction of such a design by trial-and-error leads rapidly to a solution, and there are in fact many solutions. A design with t^2 units arranged in a crossed double-blocking classification system with t blocks in each system, and with t treatments each occurring once in each block of each blocking system (so that there are t replicates of each treatment) is called a Latin square design, the name reflecting the common use of Latin letters, A, B, C, …, to represent the treatments.

6.1.2: LSD versus RCBD

RCB design takes care of heterogeneity in one direction. Thus, there is a need for experimental design which can take care of heterogeneity among EUs in two perpendicular directions. LSD is such a design which takes care of two perpendicular sources of variations among the EUs. As per the principle of blocking, blocking is to be done in perpendicular direction of heterogeneity. Thus, we need to frame blocks in two perpendicular directions independently. And as per characteristics of blocking, each block should contain each treatment once in each block.

Similarly, in terms of number of factors, there are two factors in RCBD and three in case of LSD.

All the knowledge belongs to God.

6.1.3: Merits and Demerits of LSD

LSD takes care of heterogeneity in two perpendicular directions. In the absence of any knowledge about the experimental site, it is better to have LSD. Among the three basic designs, LSD is the most efficient design, particularly toward error minimization.

However, the number of replications equals to the number of treatments; thereby an increased number of EUs are required for conduction of an experiment compared to other two basic designs, viz., CRD and RBD. The layout of the design is not so simple as was in the case of CRD or RBD and it requires square number plots.

6.2: LSD FORMULATION (ONE FACTOR+TWO CONSTRAINTS)

Suppose we want to check the yield of 4 types of seeds of corn breeds (F_1, F_2, F_3, and F_4) against the two constraints on a field: air speed across the field (varying from North to South) and sunlight across the field (varying from East to West).

Variables included in this example are:

- $Factor = seed$
- $Constraint\ 1 = sunlight$
- $constraint\ 2 = air\ speed$
- $response = corn\ yield$
- $experimental = field\ (soil)$

6.2.1: Hypothesis

Here we will have at least three hypotheses only for main effects as there is no interaction involved:

$$H_o = All\ the\ breeds\ have\ equal\ response$$
$$H_1 = At\ least\ one\ breed\ has\ different\ response\ from\ others$$

$$H_o' = Air\ speed\ across\ the\ north\ south\ field\ gradient\ have\ same\ response$$
$$H_1' = At\ least\ one\ field\ across\ the\ north\ south\ field\ with\ air\ speed\ gradient$$
$$have\ different\ response$$

$$H_o'' = Sunlight\ across\ the\ east\ west\ field\ gradient\ have\ same\ response$$
$$H_1'' = At\ least\ one\ filed\ across\ the\ east\ west\ field\ with\ sunlight\ gradient$$
$$have\ different\ response$$

6.2.2: Model

$$Corn\ yield = (Corn\ breed + air\ speed + sunlight) + Error$$

6.2.3: Layout

The layout of the LSD starts with a standard Latin square. *A standard Latin square is an arrangement in which the treatments are arranged in natural/alphabetical order or systematically.* Then in the next step, columns are randomized, and in the last step keeping the first row intact, the rest of the rows are randomized.

The constraints will be written outside the table and the factor will be written inside the table. Remember that the response will be written inside the table along with factors in brackets (which is not written here). All 4 levels (four breeds) of factor (corn seed) will be randomly allocated along both gradients (Table 6.2), but don't repeat the treatment in any block of any side:

Table 6.2: Layout of LSD (1 factor and 2 constraints).

Sunlight (north south)

All the knowledge belongs to God.

	S_1	S_3	S_2	S_4
A_1	A	B	C	D
A_3	B	D	A	C
A_2	C	A	D	B
A_4	D	C	B	A

(Air speed (East west))

Notice that the design is a square arrangement and that the five formulations (or treatments) are denoted by the Latin letters A, B, C, D, and E; hence the name Latin square. We have 16 cells here: four levels across two gradients: (4×4). In LSD, each block in perpendicular directions, i.e., each row block and column block, should contain each and every treatment once and only once.

Thus, the rows and columns actually represent two restrictions on randomization. In general, a Latin square for p factors, or a $p \times p$ Latin square, is a square containing p rows and p columns. Each of the resulting p^2 cells contains one of the p letters that corresponds to the treatments, and each letter occurs once and only once in each row and column.

6.2.4: Design

It is LSD or three-way ANOVA.

6.3: LSD FORMULATION (THREE FACTORS OR FACTORIAL LSD)

Suppose we want to check the yield of 4 types of seeds of corn breeds (A, B, C and D) with four types of fertilizers (F_1, F_2, F_3, and F_4) with four levels of water (1, 2, 3 and 4) on the same field (no variation is involved).

6.3.1: Hypothesis

Here we will have seven hypotheses: one for each main effect and one for each interaction. To avoid this lengthy task, we can write only interaction effect in interaction hypothesis in which all the factors interact (last one):

$$H_o = All\ the\ breeds\ have\ equal\ response$$
$$H_1 = Change\ in\ breed\ has\ an\ effect\ on\ yield$$

$$H_o{}' = All\ the\ fertilizers\ have\ same\ response$$
$$H_1{}' = Change\ in\ fertilizer\ has\ an\ effect\ on\ yield$$

$$H_o{}'' = All\ the\ water\ levels\ have\ same\ response$$
$$H_1{}'' = Change\ in\ water\ level\ has\ an\ effect\ on\ yield$$

$$H_o{}''' = All\ the\ interactions\ (breeds, feritlizers\ and\ water\ levels)\ have\ same\ response$$
$$H_1{}''' = There\ exist\ an\ interaction\ between\ corn\ breeds, fertilizer\ and\ water\ levels$$

All the knowledge belongs to God.

6.3.2: Model

In writing model, LSD have 7 effects in it. We cannot skip the interactions; we will have to write down all the interactions along with main effects as following:

$$Yield = (breed + air\ speed + sunlight) + (breed \times fertilizer) + (breed \times water)$$
$$+ (fertilizer \times water) + (breed \times fertilizer \times water) + error$$

As we have written down all the factor levels, that is why it is also known as **full factorial.** If we have taken only the interaction (in which we are interested), say fertilizer and water by skipping the rest of interactions, it will be called as **customized factorial**.

6.3.3: Layout

Two factors will be written outside the table and one factor will be written inside the table (Table 6.3). Remember that the response will be written inside the table along with inside-factor in brackets (which is not written here). All four levels of factors will be randomly allocated along both sides of table.

Table 6.3: Layout of factorial LSD.

		Fertilizers			
		F_1	F_3	F_2	F_4
Water levels	**2**	A	B	C	D
	4	B	D	A	C
	1	C	A	D	B
	3	D	C	B	A

6.3.4: Design

It is LSD under factorial or simply factorial design.

6.4: DESIGN RECOGNITION

6.4.1: Example 1

A veterinarian conducted an experiment to compare the yield of four varieties of deer (A, B, C and D). An enclosure of land was divided into 16 sub-enclosures (four rows and four columns) according to the sunlight and vegetation. The following design (Table 6.4) was run and the responses are given.

Table 6.4: Yield of four fish varieties.

Treatment (deer variety)						Response (yield)			
Row	Column					Column			
N	C	A	B	D	Row	E	EC	WC	W
NC	A	B	D	C	N	26	19	29	30
SC	B	D	C	A	NC	23	22	25	29
S	D	C	A	B	SC	29	20	29	27
					S	25	17	29	35

Variables included in above problem are:

All the knowledge belongs to God.

- *Factor = deer varieties*
- *Constraint* 1 = *horizontal enclosure* (*rows*)
- *Constraint* 2 = *perpendicular enclosure* (*columns*)
- *Response = yield*
- *Experimental units = enclosure*

The hypotheses will be:

$$H_1 = Change\ in\ varieties\ has\ an\ effect\ on\ yield$$

$$H_1' = Change\ in\ horizontal\ enclosure\ has\ an\ effect\ on\ yield$$

$$H_1'' = Change\ in\ perpendicular\ enclosure\ has\ an\ effetc\ on\ yield$$

It is LS design, and the data can be analyzed using three-way ANOVA.

6.4.2: Example 2

A poultry scientist wants to know the effect of light (warm white and cool white), feed (Islamabad and SB) and spacing (normal and narrow) on the weight gain of broiler chickens.

Table 6.5: weight gain of broilers.

	Light	
Feed	Warm white	Cool white
Islamabad	Narrow	Narrow
	Normal	Normal
SB	Narrow	Narrow
	Normal	Normal

Variables included are:
- *Factor* 1 = *Feed* (2 *levels*).
- *Factor* 2 = *light* (*warm white and cool white*)
- *Factor* 3 = *spacing* (*normal and narrow*).
- *Response = weight gain*
- *EUs = broilers*

The hypotheses will be:

$$H_1 = Change\ in\ the\ feed\ has\ an\ effect\ on\ weight\ gain$$

$$H_1' = Change\ in\ the\ light\ has\ an\ effect\ on\ weight\ gain$$

$$H_1'' = Change\ in\ the\ spacing\ has\ an\ effect\ on\ weight\ gain$$

$$H_1''' = There\ exists\ an\ interaction\ among\ feed, light\ and\ spacing$$

It is factorial LSD and three-way ANOVA can be used for analysis.

All the knowledge belongs to God.

CHAPTER 7: GRAECO LATIN SQUARE DESIGN

7.1: INTRODUCTION

Randomized Complete Blocks allow us to control one extraneous source of variability in our units, and Latin Squares allow us to control two sources. The Latin Square design can be extended to control for three sources of extraneous variability; this is the Graeco-Latin Square. Similarly, for four or more sources of variability, we use Latin Hyper-Squares.

GLSD allows investigation of four factors (rows, columns, Latin letters, and Greek letters), each at p levels in only p^2 runs. It is so-called four-way ANOVA. Beyond Graeco-LSD (GLSD), there can be more than 4 factors involved in problems with solution.

Like RCBD, if 4 factors have same levels it will be called GLSD otherwise it will be known as 4-way ANOVA and not the GLSD. Remember that every GLSD is a 4-way ANOVA but the inverse is not true.

7.2: GLSD FORMULATION (SYMMETRICAL)

Suppose we want to check the yield of three types of corn seeds (C1, C2 and C3) with three fertilizers (F1, F2 and F3), under three amounts of sunlight (S1, S2 and S3) at three water levels (W1, W2 and W3).

7.2.1: Hypothesis

There will be 15 hypotheses: four hypotheses for each main effects and 11 for interaction effects, but to avoid lengthy task we can write only the interaction which involves all the factors:

$$H_o = All\ the\ corn\ breeds\ have\ equal\ response$$
$$H_1 = At\ least\ one\ corn\ breed\ has\ different\ response\ from\ others$$

$$H_{o1} = All\ the\ fertilizers\ have\ equal\ response$$
$$H_{11} = At\ least\ one\ type\ of\ fertilizer\ has\ different\ response\ from\ others$$

$$H_{o2} = All\ the\ types\ of\ fertilizers\ have\ equal\ response$$
$$H_{12} = At\ least\ one\ type\ of\ fertilizer\ has\ different\ response\ from\ others$$

$$H_{o3} = All\ the\ amounts\ of\ sunlight\ have\ equal\ response$$
$$H_{13} = At\ least\ one\ amount\ of\ sunlight\ has\ different\ response\ from\ others$$

$$H_{o4} = All\ the\ water\ levels\ have\ equal\ response$$
$$H_{14} = At\ least\ one\ water\ level\ has\ different\ response\ from\ others$$

$$H_{o5} = All\ the\ interactions\ between\ breeds + fertilizers + sunligh + \\ water\ levels\ have\ equal\ response$$
$$H_{15} = At\ least\ one\ interaction\ has\ different\ response\ from\ other\ interactions$$

7.2.2: Model

$$Yield = (B + F + S + W) + (B \times F) + (B \times S) + (B \times W) + (F \times S) + (F \times W) \\ + (S \times W) + (B \times F \times S) + (B \times S \times W) + (F \times S \times W) \\ + (B \times F \times S \times W)$$

7.2.3: Layout

Each entry in the Table 7.1 has one Latin letter and one Greek letter. Latin letters correspond to treatments, as in a Latin Square, and Greek letters correspond to the third

All the knowledge belongs to God.

blocking factor. The Latin letters occur once in each row and column (they form a Latin Square), and the Greek letters occur once in each row and column (they also form a Latin Square). In addition, each Latin letter occurs once with each Greek letter. Here is a 4×4 Graeco-Latin Square:

Table 7.1: Typical GLSD layout.

A α	B γ	C δ	D β
B β	A δ	D γ	C α
C γ	D α	A β	B δ
D δ	C β	B α	A γ

We can only draw table for 3 factors, but we have 4 factors in GLSD, so we will have to fix one factor and draw the table under the levels of factor which we have fixed already like a hierarchical model (Table 7.2). Let's say we fix corn breeds (C) as:

Table 7.2: Layout of symmetrical GLSD.

Corn breeds		
C1	**C2**	**C3**

C1

Fertilizers			
	F1	F2	F3
	S1	W1	W2	W3
Sunlight	S2	W2	W3	W1
	S3	W3	W1	W2

C2

Fertilizers			
	F2	F1	F3
	S3	W1	W2	W3
Sunlight	S1	W2	W3	W1
	S2	W3	W1	W2

C3

Fertilizers			
	F1	F3	F2
	S3	W1	W2	W3
Sunlight	S2	W2	W3	W1
	S1	W3	W1	W2

The response is written inside brackets with the fourth factor in table cells. As we have written down all the levels of all the factors, that is why it is also known as **full factorial**.

7.2.4: Design

It is the GLSD or four-way ANOVA (as all the levels are equal).

7.3: GLSD FORMULATION (ASYMMETRICAL)

Suppose we want to check the yield of three types of corn seeds (C1, C2 and C3) with two fertilizers (F1 and F2), under two amounts of sunlight (S1 and S2) at five water levels (W1, W2, W3, W4 and W5).

7.3.1: Hypothesis

Hypotheses will be the same as previous GLSD example:

$$H_o = All\ the\ corn\ breeds\ have\ equal\ response$$
$$H_1 = At\ least\ one\ corn\ breed\ has\ different\ response\ from\ others$$

$$H_{o1} = All\ the\ fertilizers\ have\ equal\ response$$
$$H_{11} = At\ least\ one\ type\ of\ fertilizer\ has\ different\ response\ from\ others$$

$$H_{o2} = All\ the\ types\ of\ fertilizers\ have\ equal\ response$$
$$H_{12} = At\ least\ one\ type\ of\ fertilizer\ has\ different\ response\ from\ others$$

All the knowledge belongs to God.

$$H_{o3} = All\ the\ amounts\ of\ sunlight\ have\ equal\ response$$
$$H_{13} = At\ least\ one\ amount\ of\ sunlight\ has\ different\ response\ from\ others$$

$$H_{o4} = All\ the\ water\ levels\ have\ equal\ response$$
$$H_{14} = At\ least\ one\ water\ level\ has\ different\ response\ from\ others$$

$$H_{o5} = All\ the\ interactions\ between\ breeds + fertilizers + sunlight +$$
$$water\ levels\ have\ equal\ response$$
$$H_{15} = At\ least\ one\ interaction\ has\ different\ response\ from\ other\ interactions$$

7.3.2: Model

Model will also be the same as above example of GLSD:

$$Yield = (B + F + S + W) + (B \times F) + (B \times S) + (B \times W) + (F \times S) + (F \times W)$$
$$+ (S \times W) + (B \times F \times S) + (B \times S \times W) + (F \times S \times W)$$
$$+ (B \times F \times S \times W)$$

7.3.3: Layout

Layout will be different as levels of factors have now changed (Table 7.3). For the ease of drawing table, it is advised to fix that factor which has minimum levels otherwise the sheet will have to be used in landscape orientation. Let's say we fix fertilizers (F):

Table 7.3: Layout of asymmetrical GLSD.

Fertilizers								
F1							**F2**	
Corn breeds					**Corn breeds**			
		C1	C2	C3			C1	C1	C1

		Corn breeds					Corn breeds		
		C1	C2	C3			C1	C1	C1
Sunlight	S1	W1	W2	W3	**Sunlight**	S3	W1	W2	W4
	S2	W2	W3	W1		S2	W5	W3	W5

The response is written inside brackets with the fourth factor in table cells. Note that the water has 5 levels, but we cannot write all levels inside table, so we can skip the others after randomly choosing the required levels. This is known as fractional factorial. Remember that the levels can be repeated while drawing the table of GLSD but not in case of LSD.

7.3.4: Design

It is GLSD but four-way ANOVA can be used (as all the levels of factors are not equal).

All the knowledge belongs to God.

CHAPTER 8: FACTORIAL EXPERIMENTS

8.1: FACTORIAL DESIGNS

It is basically a **treatment design** unlike the other three basic designs which are error-control designs. *Factorial treatment structure exists when the treatments are the combinations of the levels of two or more factors.* We call these combination of treatments factor-level combinations or factorial combinations to emphasize that each treatment is a combination of one level of each of the factors. These designs are also known as **between-groups factorial designs.**

Factorial experiments are such a mechanism in which more than one group (factor) of treatments can be accommodated in one experiment, and from the experiment, not only the best treatment in each group of treatments could be identified but also the interaction effects among the treatments in different groups could also be estimated.

8.1.1: Main Effect and Interaction

The major concepts of this factorial analysis are main effects and interaction effects. *The main effect of a factor is the effect of the factor concerned irrespective of the levels of other factors.* The main effect of factor A in any example (containing two factors) describes variation due solely to the level of factor A, and the main effect of factor B describes variation due solely to the level of factor B in that experiment.

The interaction effects are the effects of one factor with the change in levels of the other factor and vice versa. When the factors are independent of one another, one would expect the same effect of one factor at various levels of the other factors resulting in the zero-interaction effect. Thus, interaction effects come into picture when the factors are not independent. In other words, variation that arises from changing rows and columns simultaneously, we call such variation interaction between factors A and B. An interaction is a deviation from additivity. For example, if the effect of going from dose 1 to dose 2 changes from drug 2 to drug 3, then there is an interaction between drug and dose. Similarly, if the interaction of drug and dose is different in morning and evening applications, then there is a three-factor interaction between drug, dose, and time.

Here is another way to think about main effect and interaction. The main effect of rows tells us how the response changes when we move from one row to another, averaged across all columns. The main effect of columns tells us how the response changes when we move from one column to another, averaged across all rows. The interaction tells us how the change in response depends on columns when moving between rows, or how the change in response depends on rows when moving between columns.

8.1.2: Need of Factorial Design?

Analysis of factorially structured data should be more than just an enumeration of which main effects and interactions are significant. For example, reporting that factor B only affects the response at the high level of factor A is more informative than reporting that factors A and B have significant main effects and interactions.

8.1.3: Factorial Symmetry

Factorial experiment is either **symmetrical** (the same levels for all the factor) also known as **pure factorial experiment** or **asymmetrical** (different levels for different factors) also known as **mixed factorial experiment**, respectively: e.g., a two-factor factorial experiment with five breeds and five different diets is a symmetrical factorial experiment. On the other hand, a two-factor factorial experiment with five varieties and three diets is an asymmetrical factorial experiment.

All the knowledge belongs to God.

Generally, a symmetrical factorial experiment with "n" factors each at "m" levels is denoted as $\mathbf{m^n}$ **factorial experiment**. Thus, a 2^3 factorial experiment is a symmetrical factorial experiment with 3 factors each at 2 levels. A two-series factorial design is one in which all the factors have just two levels. For k factors, we call this a $\mathbf{2^k}$ **design**, because there are 2^k different factor-level combinations. Similarly, a design with k factors, each with three k levels, is a three-series design and denoted by 3^k. Two-series designs are somewhat special, because they are the smallest designs with k factors. Remember that the model and analysis of a multi-way factorial are like those of a two-way factorial.

An asymmetrical two-factorial experiment with p levels for the first factor and q levels for the second factor is denoted as $\boldsymbol{p \times q}$ **factorial experiment**. Thus a 2×3 asymmetrical factorial experiment means there are two factors in the experiment with the first factor having 2 levels and the second factor having 3 levels.

8.1.4: Types of Factorial Experiments

Factorial experiments are of different types. Depending upon the number of factors included in the experiment, a factorial experiment is a two-factor factorial experiment (when two sets of treatments, i.e., two factors), three-factor factorial experiment (when three sets of treatments, i.e., three factors) ...., or p-factor factorial experiment (when "p" sets of treatments, i.e., p number of factors). Depending upon the number of factors involved in the factorial experiment, the interaction effects would be the first-order interaction, second order interaction, third-order interaction, and so on when the number of factors in the experiment is 2, 3, 4, and so on. When two factors are involved in a factorial experiment, then we are concerned about the interaction effects of two factors, known as the **first-order interaction.** But when more than two factors are involved, then the interaction will be pairwise as well as overall. That means for a three-factor factorial experiment we shall have two-factor interaction, i.e., **first-order interaction,** as well as the three-factor interaction, i.e., **second-order interaction.**

8.1.5: OFAT versus Factorial Experiment

It is easiest to see the advantages of factorial treatment structure by comparing it to OFAT wherein we only vary the levels of a single factor. When the factors interact, factorial experiments can estimate the interaction. OFAT experiments cannot estimate interaction. Use of OFAT experiments in the presence of interaction can lead to serious misunderstanding of how the response varies as a function of the factors (this is the main reason why do postgraduate papers fail in field). Even, when the factors do not interact, factorial experiments are more efficient than OFAT experiments, in that the units can be used to assess the (main) effects for both factors. Units in OFAT experiment can only be used to assess the effects of one factor. There are thus two times when you should use factorial treatment structure—when your factors interact, and when your factors do not interact. Factorial structure is a win, whether or not we have interaction.

8.1.6: Fractional Factorial Design

A fractional factorial design was introduced first by **Finney**. In factorial experiments the fact that the number of treatment combinations increases rapidly as the number of factors and/or levels increases. One way out of this dilemma is to consider only a *subset of all possible treatment combinations, a so-called fractional factorial.* If it uses half of the run, it will be called as **one-half fraction**.

Cost in money, time and effort are always prohibitive in large factorial, except in small factorial experiment. To avoid such constraints, researcher may use fractional factorial design, in which only subset of complete runs is used or in which certain portion of full factorial is used (say

All the knowledge belongs to God.

only 8 out of 128 cells) rather all combinations (128 cells). But there is a price to pay for it, as the probability that 1 out of 8 combinations would be optimum is 8 out of 128 (6.25%).

8.1.7: Merits and Demerits of Factorial Experiments

When the factors interact, factorial experiments can estimate the interaction. But when the factors do not interact, factorial experiments are more efficient than OFAT experiments, in that the units can be used to assess the (main) effects for both factors. Factorial experiments can accommodate more than one set of treatments (factors) in one experiment and are resource and time saving as well. The required minimum degrees of freedom for error components in the ANOVA can easily be achieved in factorial experiments compared to single factorial experiments.

However, with the increase in number of factors or the levels of the factors or both the number and levels of factors are more, the number of treatment combinations will be more, resulting in the requirement of bigger experimental area and bigger block size. As the block size increases, it is very difficult under field conditions to maintain homogeneity among the plots within the block. Thus, there is a possibility of increasing the experimental error vis-a`-vis decrease in the precision of experiment. The layout of the factorial experiment is comparatively difficult than simple experiments. Statistical procedure and calculation of factorial experiments are more complicated than the single factor experiments. With the increase in the number of factors or the levels of the factor or both, the number of effects, including the interaction effects, also increases. Sometimes it becomes very difficult to explain the information from interactions, particularly the higher-order interaction effects. The risk is high. Failure in one experiment may result in greater loss of information compared to single-factor simple experiments.

8.2: FACTORIAL DESIGN FORMULATION

Suppose we have same example as mentioned in RCBD, but now we have two factors: feed choices (A, B, C, D and E) with different vitamins (1 and 2) on broilers for yield. We will find the best interaction which can produce the maximum yield.

Variables included in this problem are:
- $Factor\ 1 = feed\ (4\ levels)$
- $Factor\ 2 = vitamin\ (2\ levels)$
- $Response = yield\ (weight)$
- $Experimental\ units = broiler$

8.2.1: Hypothesis

There will be 3 hypotheses in this factorial design: one for each factor (main effects) and one for interaction effect of these factors:

$$H_o = All\ the\ feed\ choices\ have\ the\ same\ response$$
$$H_1 = At\ least\ one\ feed\ has\ different\ response\ to\ others$$

$$H'_o = All\ vitamins\ have\ equal\ response$$
$$H'_1 = At\ least\ one\ vitamin\ has\ different\ repsonse\ than\ others$$

$$H''_1 = All\ the\ interactions\ of\ feed\ choices\ with\ vitamins\ have\ equal\ response.$$
$$H''_1 = There\ exist\ an\ interaction\ betweem\ feed\ and\ vitamins.$$

Notice the use of word 'with' in last hypothesis showing the interaction involvement, making it factorial design. And this is the reason that only the significance of

All the knowledge belongs to God.

interaction is taken as proof of the test hypothesis and not the significance of the 'main effect'.

8.2.2: Model

Main effect and interaction are written separately in predicate structure while writing factorial model:

$$Weight\ gain = (Feed + vitamins) + (Feed \times vitamins) + Error$$

8.2.3: Layout

It is an 5×2 asymmetrical factorial design and layout will be as following:

Table 8.1: Response under vitamins and feeds interaction.

		Vitamins	
		1	2
Feed	**Feed A**	$Feed\ A \times 1$	$Feed\ A \times 2$
	Feed B	$Feed\ B \times 1$	$Feed\ B \times 2$
	Feed C	$Feed\ C \times 1$	$Feed\ C \times 2$
	Feed D	$Feed\ D \times 1$	$Feed\ D \times 2$
	Feed E	$Feed\ E \times 1$	$Feed\ E \times 2$

8.2.4: Design Name

Factorial design and factorial ANOVA (two-way) for analysis.

8.3: DESIGN RECOGNITION

8.3.1: Example 1

A poultry physiologist investigated the effect of mechanical stress on the growth of broilers. They were randomly allocated to four treatment groups of 13 broilers each. Broilers in two groups were stressed by high temperature for 20 minutes twice daily, while two control groups were not stressed. After 14 days of growth, the broilers were slaughtered, and the total weight gain of each broiler was taken. The results are given in Table 8.2. Recognize the design and analyze data.

Table 8.2: Broiler growth under stress and control conditions.

Low		Moderate	
Control	**Stress**	**Control**	**Stress**
264	235	314	283
200	188	320	312
225	195	310	291
268	205	340	259
215	212	299	216
241	214	268	201
232	182	345	267
256	215	271	326
229	272	285	241
288	163	309	291
253	230	337	269
288	255	282	282
230	202	273	257

All the knowledge belongs to God.

Variables included in this example are:
- $Factor\ 1 = stress$
- $Factor\ 2 = light$
- $Response = broiler\ growth$
- $Experimental\ units = broilers$

The hypotheses of this example will be:
$$H_1 = Stress\ has\ an\ effect\ on\ broiler\ growth$$

$$H_1' = Light\ has\ an\ effect\ on\ broiler\ growth$$

$$H_1'' = There\ exist\ an\ interaction\ among\ stress\ and\ light\ on\ broiler\ growth.$$

The model in this example will be:
$$Growth = (stress + light) + (stress \times light) + Error$$

It is factorial RCBD, and factorial ANOVA can be used in analysis. Because two factors (stress and light) are involved. Notice in Table 8.2 that we have the same number of responses in every factor-level combination known as '**balance**'. Balance turns out to be important for the standard analysis of factorial responses replication.

8.3.2: Example 2

Data collected on the growth (times of initial weight) and genotype in pure bred chicken is given Table 8.3. Because it is not known whether sex also affected growth, we separated the broilers by sex. Each genotype contains some males and some females, and each sex contains all three genotypes.

Table 8.3: Growth due to gender and genotype.

Genotype	FF				FS				SS			
Male	1.8	2.2	4.9	3.4	2.3	2.9	3.1	2.6	2.8	3.4	4.2	3.1
Female	2.8	4.2	2.8	4.1	3.5	4.5	3	1.9	3.6	3	3.5	2.6

Variables included in this example are:
- $Factor\ 1 = genotype\ (3\ levels)$
- $Factor\ 2 = gender\ (2\ levels)$
- $Response = growth$
- $Experimental\ units = broilers$

The hypotheses will be:
$$H_1 = At\ least\ one\ genotype\ has\ an\ effect\ on\ growth$$

$$H_1' = At\ least\ one\ gender\ has\ an\ effect\ on\ growth$$

$$H_1'' = There\ exist\ an\ interaction\ between\ gender\ and\ genotype$$

The model of this example will be:
$$Growth = (genotype + gender) + (genotype \times gender) + Error$$

All the knowledge belongs to God.

It is factorial design or factorial ANOVA.

8.3.3: Example 3

An investigator believes that students who rank high in certain kinds of motives will behave differently in gambling situations. To investigate this hypothesis, the investigator randomly samples 50 students high in affiliation motivation, 50 students high in achievement motivation, and 50 students high in power motivation. The students are asked to play the game of roulette, and a record is kept of the bets they make. The data are then grouped into the number of subjects with each kind of motivation who make bets involving low, medium, and high risk (Table 8.4). Low risk means they make bets involving low odds (even money or less), medium risk involves bets of medium odds (from 2 to 1 to 5 to 1), and high risk involves playing long shots (from 17 to 1 to 35 to 1). Analyze.

Table 8.4: Bets count under kind of motives and risk.

	Kind of Motive		
	Affiliation	**Achievement**	**Power**
Low risk	26	13	9
Medium risk	16	27	14
High risk	8	10	27

Using α= 0.05, is there a relationship between these different kinds of motives and gambling behavior? How do the groups differ?

Variables included are:

- $Factor\ 1 =\ kind\ of\ motives\ (3\ levels)$
- $Factor\ 2 =\ risk\ (3\ levels)$
- $Response =\ bets$
- $EUs =\ students$

The hypotheses will be:

$$H_1 = At\ least\ one\ of\ the\ motive\ has\ different\ bets$$

$$H_1' = At\ least\ one\ of\ the\ risk\ level\ has\ different\ bets$$

$$H_1'' = There\ exist\ an\ interaction\ between\ risk\ and\ motive\ regarding\ bets$$

The model so will be:

$$Growth = (motives + risk) + (motives \times risk) + Error$$

It is factorial RCBD or factorial ANOVA.

All the knowledge belongs to God.

CHAPTER 9: INTRODUCTION TO SPSS

9.1: INTRODUCTION TO SPSS

We will explore the key windows in SPSS (the editor, the data viewer and syntax editor) and also look at how to create variables, enter data and adjust the properties of your variables. SPSS mainly uses two windows (Figure 9.1): **The Data Editor** (this is where you input your data and carry out statistical functions) and **The Viewer** (this is where the results of any analysis appear). There are several additional windows that can be activated such as the SPSS **Syntax Editor**, which allows you to enter SPSS commands manually (rather than using the window-based menus). For many people, the syntax window is redundant because you can carry out most analyses by clicking merrily with your mouse. However, there are various additional functions that can be accessed using syntax and, in many situations, it can save time.
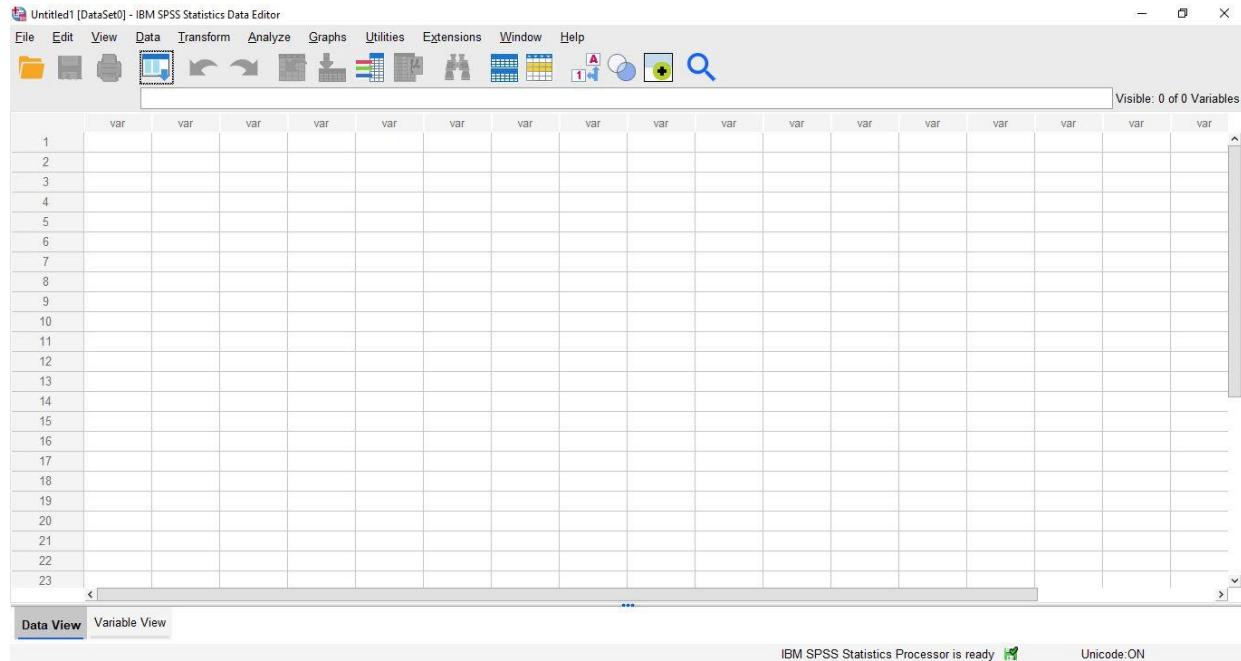
Figure 9.1: SPSS main window.

9.1.1: Data Editor

The main SPSS window includes a data editor for entering data. This window is where most of the action happens. At the top of this screen is a menu bar similar to the ones you might have seen in other programs. Figure 9.1 shows this menu bar and the data editor. There are several menus at the top of the screen that can be activated by using the computer mouse.

The data editor has two views: the **Data View** and the **Variable View** (Figure 9.2). The data view is for entering data into the data editor, and the variable view allows us to define various characteristics of the variables within the data editor. At the bottom of the data editor, you should notice that there are two tabs labelled 'Data View' and 'Variable View' and all we do to switch between these two views is click on these tabs (the highlighted tab tells you which view you're in, although it will be obvious).

All the knowledge belongs to God.
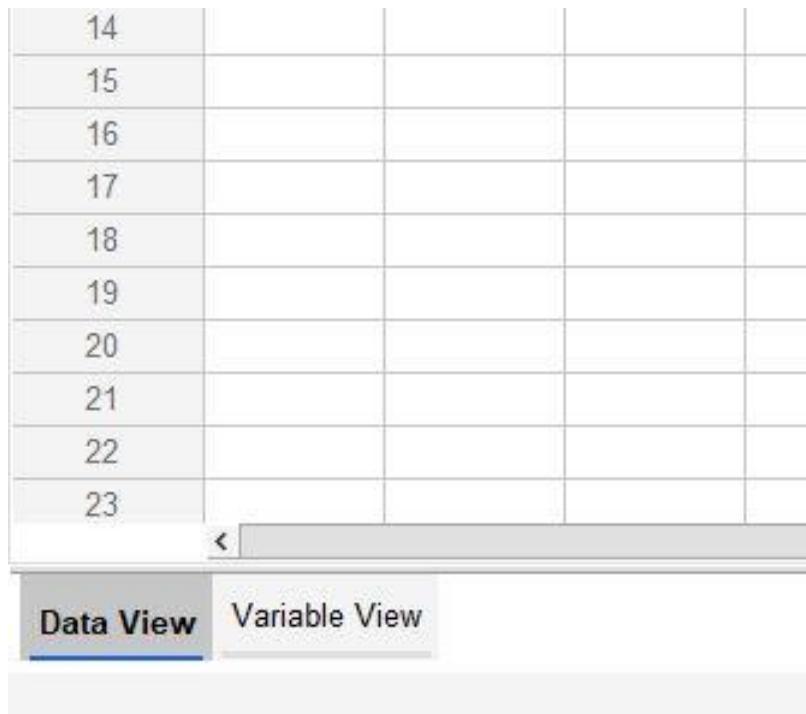
Figure 9.2: Data View and Variable View tabs in SPSS.

9.1.1.1   Variable View

When you first load SPSS it will provide a blank data editor with the title Untitled1 (this of course is daft because once it has been given the title 'untitled' it ceases to be untitled!). When inputting a new set of data, you must input your data in a logical way. The SPSS **Data Editor** is arranged such that each row represents data from one entity while each column represents a variable (Figure 9.3). There is no discrimination between independent and dependent variables: both types should be placed in a separate column. The key point is that each row represents one entity's data (be that entity a human, mouse, tulip, business, or water sample). Therefore, any information about that case should be entered across the data editor.

From the **Variable View** screen, SPSS allows you to create and edit all the variables in your data file. Each column represents some property of a variable, and each row represents a variable. All variables must be given a name. To do that, click on the first empty cell in the **Name** column and type a valid SPSS variable name. The program will then fill in default values for most of the other properties.
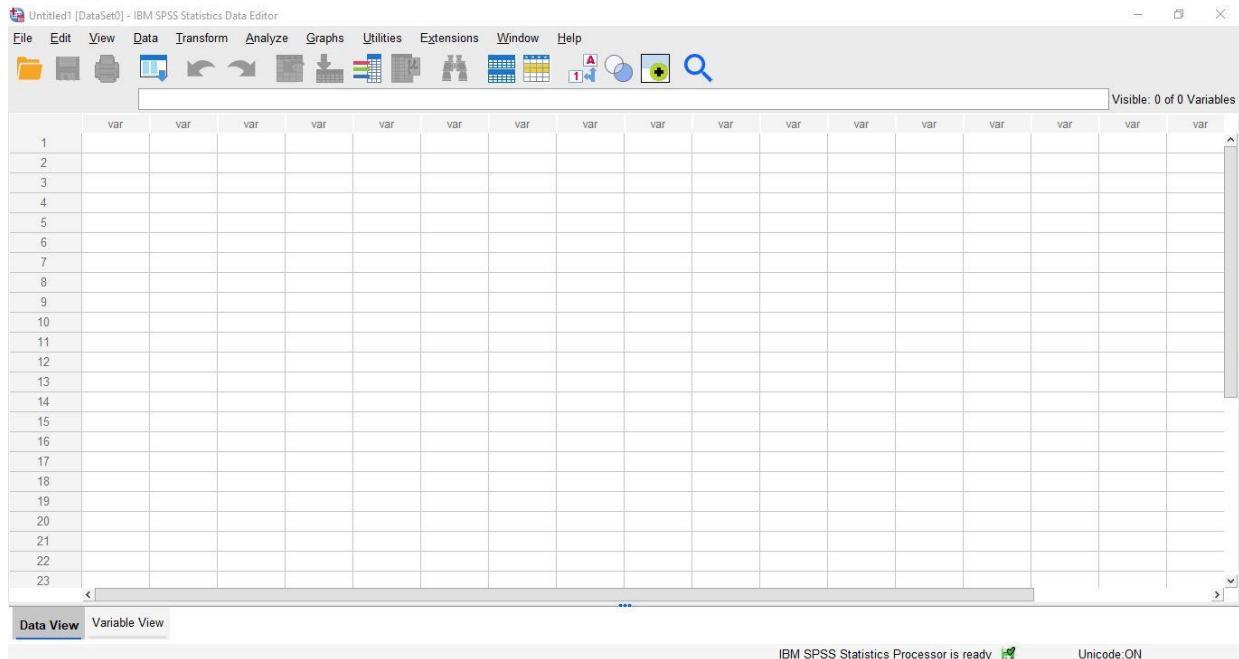
Figure 9.3: New file look of SPSS Data View window.

The data editor is made up of lots of cells, which are just boxes in which data values can be placed. When a cell is active it becomes highlighted in blue. You can move around the data editor, from cell to cell, using the arrow keys → ← etc., (found on the right of the keyboard) or by clicking the mouse on the cell that you wish to activate. To enter a number into the data editor simply move to the cell in which you want to place the data value, type the value, then press the appropriate arrow button for the direction in which you wish to move. So, to enter a row of data, move to the far left of the row, type the value and then press → (this process inputs the value and then moves you into the next cell on the right).

The first step in entering your data is to create some variables using the **Variable View** of the data editor, and then to input your data using the **Data View** of the data editor. We'll go through these two steps by working through an example.

Before we input any data into the data editor, we need to create the variables. To create variables, we use the **Variable View** of the data editor (Figure 9.4). To access this view, click on the **Variable View** tab at the bottom of the data editor; the contents of the window will change.

All the knowledge belongs to God.
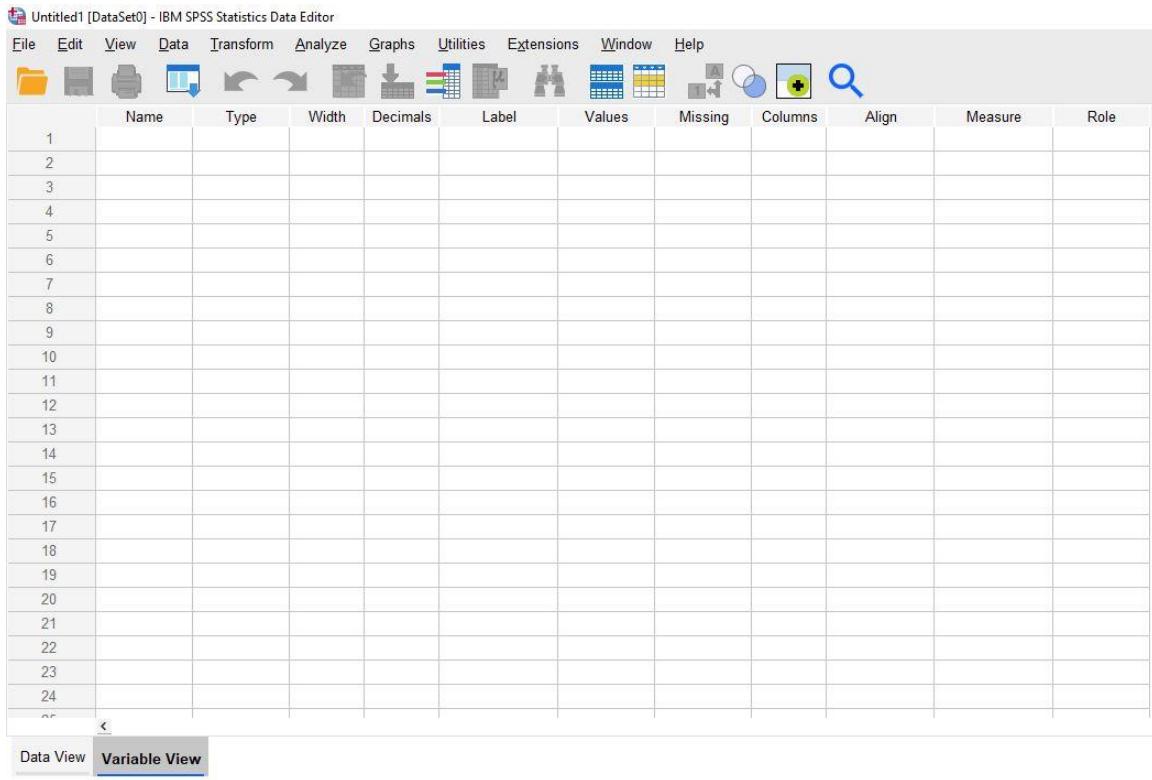
Figure 9.4: Variable View tab in SPSS.

Every row of the variable view represents a variable, and you set characteristics of a particular variable by entering information into the labelled columns (Figure 9.5). You can change various characteristics of the variable by entering information into the following columns (play around and you'll get the hang of it):
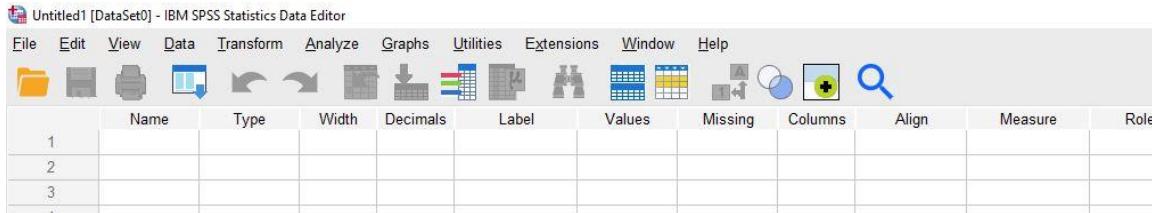
Figure 9.5: Variable View options in SPSS.

9.1.1.1.1 Name

You can enter a name in this column for each variable. This name will appear at the top of the corresponding column in the data view and helps you to identify variables in the data view (Figure 9.6). You can write what you like, but there are certain symbols you can't use (mainly symbols that have other uses in SPSS such as +, -, $, &), and you can't use spaces. (Many people use a 'hard' space in variable names, which replaces the space with an underscore, for example, i.e., Diet_type instead of Diet type.) If you use a character that SPSS doesn't like you'll get an error message saying that the variable name is invalid when you click on a different cell or try to move off the cell using the arrow keys.

All the knowledge belongs to God.

Figure 9.6: Name editing in Variable View (SPSS).

After the entry in **Variable View**, the **Data View** will look like as following (Figure 9.7):

Figure 9.7: Name option in Data View (SPSS).

9.1.1.1.2 Type

You can have different types of data. Mostly you will use numeric variables (which just means that the variable contains numbers- SPSS assumes this data type). You will come across string variables as well, which consist of strings of letters. If you wanted to type in people's names, for example, you would need to change the variable **Type** to be string rather than numeric (Figure 9.8). You can also have currency variables (i.e., £s, $s, euro) and date variables (e.g., 21-06-1973).
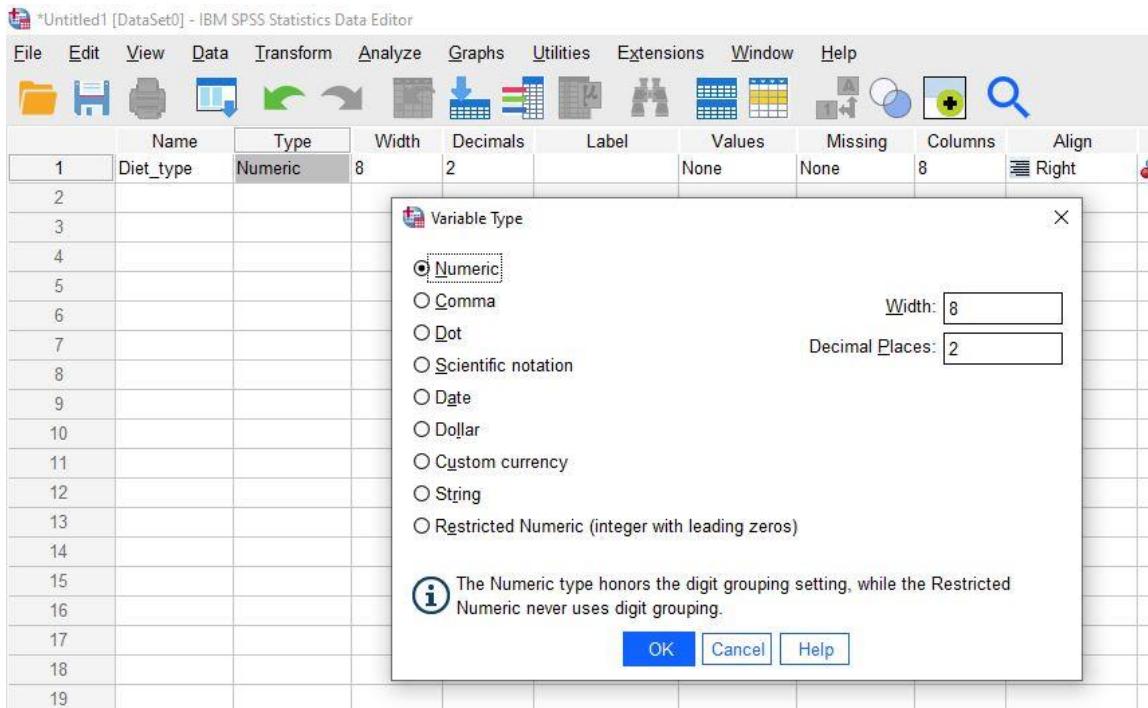
Figure 9.8: Type editing in Variable View (SPSS).

9.1.1.1.3 Width

By default, when a new variable is created, SPSS sets it up to be numeric and to store 8 digits (Figure 9.9), but you can change this value by typing a new number in this **Width** column in the dialog box. Normally 8 digits is fine, but if you are doing calculations that need to be particularly precise you could make this value bigger.

Figure 9.9: Width editing in Variable View (SPSS).

9.1.1.1.4 Decimals

Another default setting is to have 2 decimal places displayed in **Decimals** column (Figure 9.10). (You'll notice that if you don't change this option then when you type in whole numbers to the data editor SPSS adds a decimal place with two zeros after it – this can be disconcerting initially!) If you want to change the number of decimal places for a given variable then replace the 2 with a new value or increase or decrease the values using, when in the cell that you want to change.
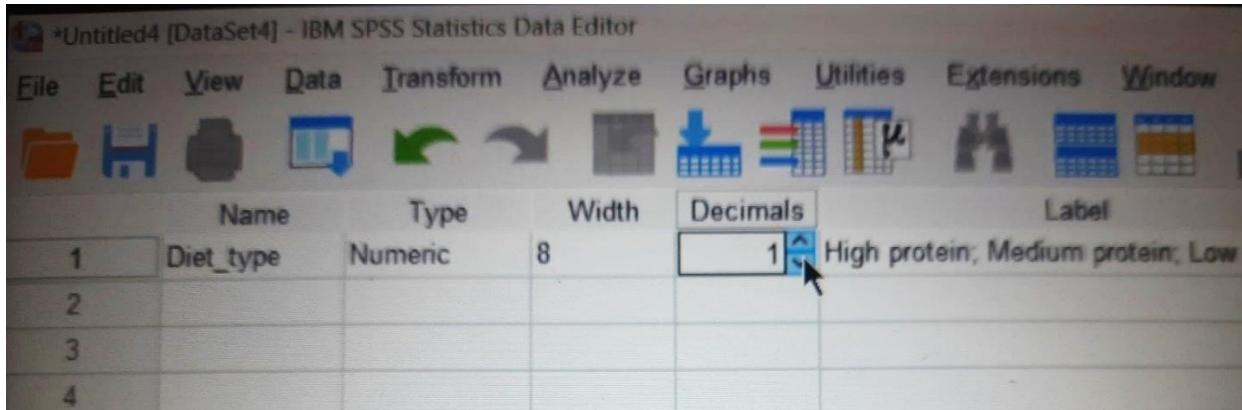
All the knowledge belongs to God.

Figure 9.10: Decimal editing in Variable View (SPSS).

9.1.1.1.5 Label

One useful function of SPSS is the ability to define variable and value labels using **Label** column (Figure 9.11). Variable labels allow you to associate a description with each variable.

The name of the variable has some restrictions on characters, and you also wouldn't want to use long names at the top of your columns (they become hard to read). Therefore, you can write a longer variable description in this column. For example, you can enter diet types (high protein; medium protein; low protein) into it for better explanation.

Figure 9.11: Label editing in Variable View (SPSS).

9.1.1.1.6 Values

Value labels allow you to associate a description with each value of a variable (Figure 9.12). This column is for assigning numbers to represent groups of people. For instance, for most procedures, SPSS requires numerical values. Thus, for data such as the day of the class (i.e., Mon/Wed/Fri and Tues/Thurs), we need to first code the values as numbers. We can assign the number 1 to Mon/Wed/Fri and the number 2 to Tues/Thurs. To help us keep track of the numbers we have assigned to the values, we use Value labels.

For example, we can code the three types of diets (low, medium and high protein diet) here for detailed output like following (Figure 1.1):
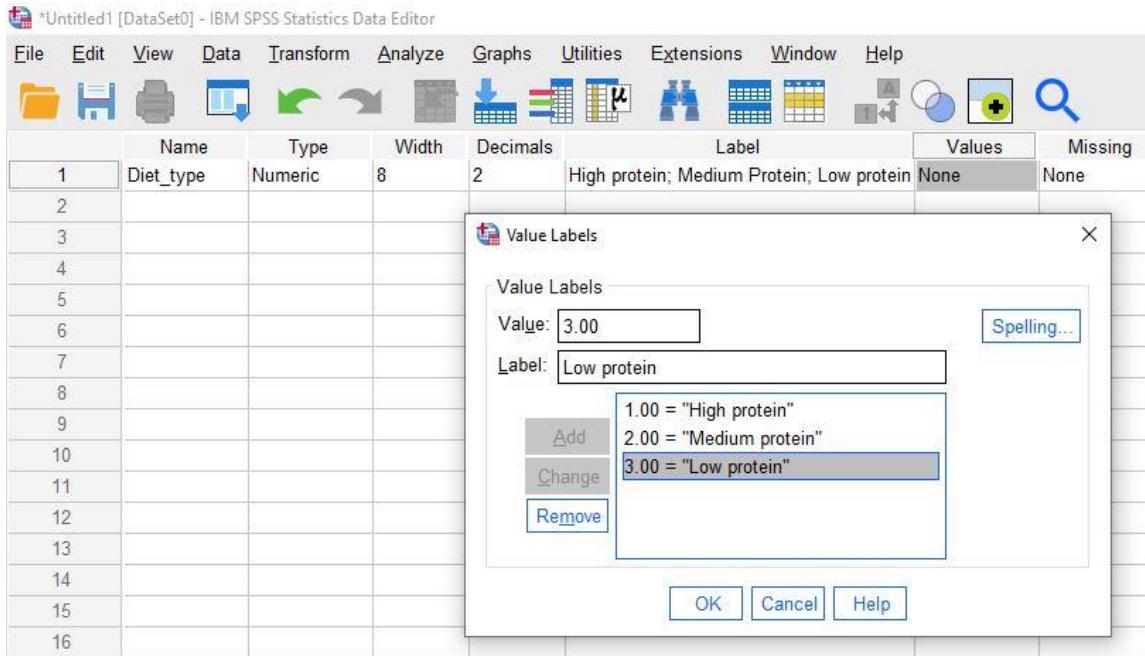
Figure 9.12: Label editing in Variable View (SPSS).

When you enter a value label, you must click **Add** after each entry. This will move the value and its associated label into the bottom section of the window. When all labels have been added, click **OK** to return to the **Variable View** window.

9.1.1.1.7 Missing

This column is for assigning numbers to missing data (Figure 9.13).
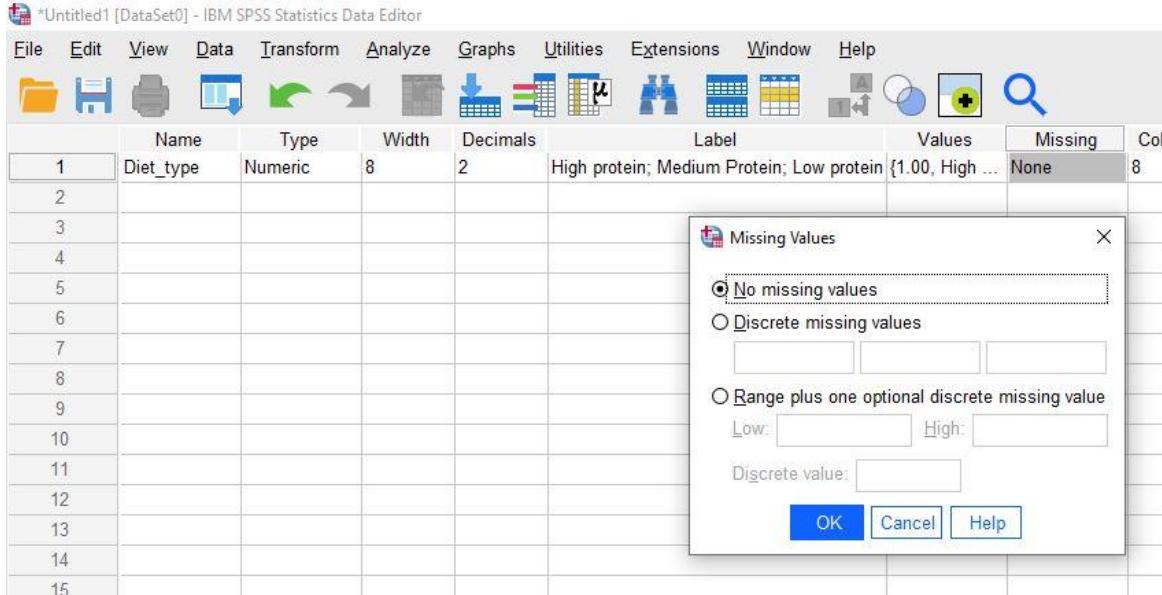
Figure 9.13: Missing tab editing in Variable View (SPSS).

Although as researchers we strive to collect complete sets of data, it is often the case that we have missing data. Missing data can occur for a variety of reasons: in long questionnaires participants accidentally (or, depending on how paranoid you're feeling,

All the knowledge belongs to God.

deliberately just to piss you off) miss out questions; in experimental procedures mechanical faults can lead to a datum not being recorded; and in research on delicate topics (e.g., sexual behavior) participants may exert their right not to answer a question. However, just because we have missed out on some data for a participant doesn't mean that we have to ignore the data we do have (although it sometimes creates statistical difficulties). Nevertheless, we do need to tell SPSS that a value is missing for a particular case.

The principle behind missing values (Table 9.1) is quite similar to that of coding variables in that we choose a numeric value to represent the missing data point. This value tells SPSS that there is no recorded value for a participant for a certain variable. The computer then ignores that cell of the data editor (it does not use the value you select in the analysis). You need to be careful that the chosen code doesn't correspond to any naturally occurring data value. For example, if we tell the computer to regard the value 4 as a missing value and several participants genuinely scored 4, then the computer will treat their data as missing when, in reality, they are not.

Table 9.1: Missing data sample.

Q1	Q2	Q3
2	2	4
3	1	
4	3	7
2		
1	2	3

If you have missing data in your dataset, leave that cell blank. In the example shown above, the fourth subject did not complete Question 2 (q2). Note that the total score (which is calculated from both questions) is also blank because of the missing data for Question 2. SPSS represents missing data in the data window with a period (although you should not enter a period—just leave it blank).

9.1.1.1.8 Column

Enter a number into **Columns** to determine the width of the column that is how many characters are displayed in the column (Figure 9.14). (This differs from **Width**, which determines the width of the variable itself – you could have a variable of 10 characters but by setting the column width to 8 you would only see 8 of the 10 characters of the variable in the data editor.) It can be useful to increase the column width if you have a string variable that exceeds 8 characters, or a coding variable with value labels that exceed 8 characters.
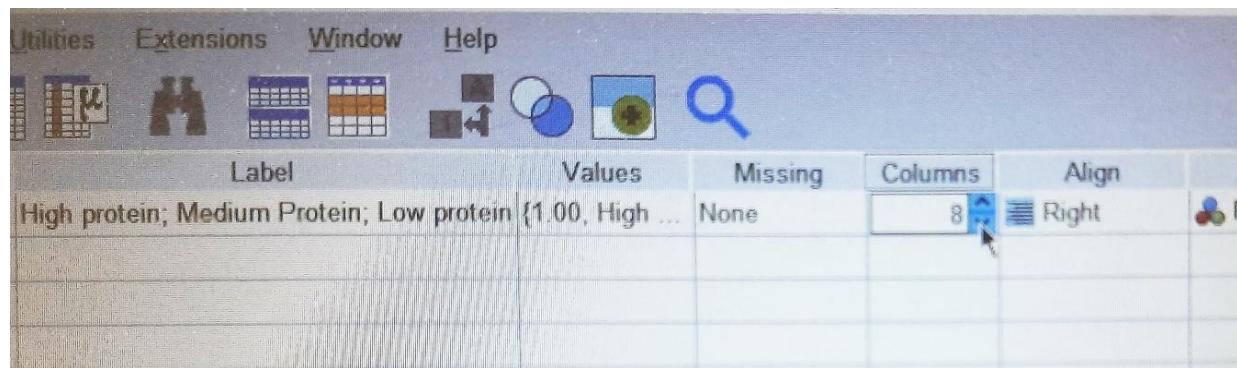
Figure 9.14: Columns tab editing in Variable View (SPSS).

All the knowledge belongs to God.

9.1.1.1.9 Align

You can use this column to select the alignment of the data in the corresponding column of the data editor. You can choose to align the data to the **Left**, **Right** or **Center** (Figure 9.15).
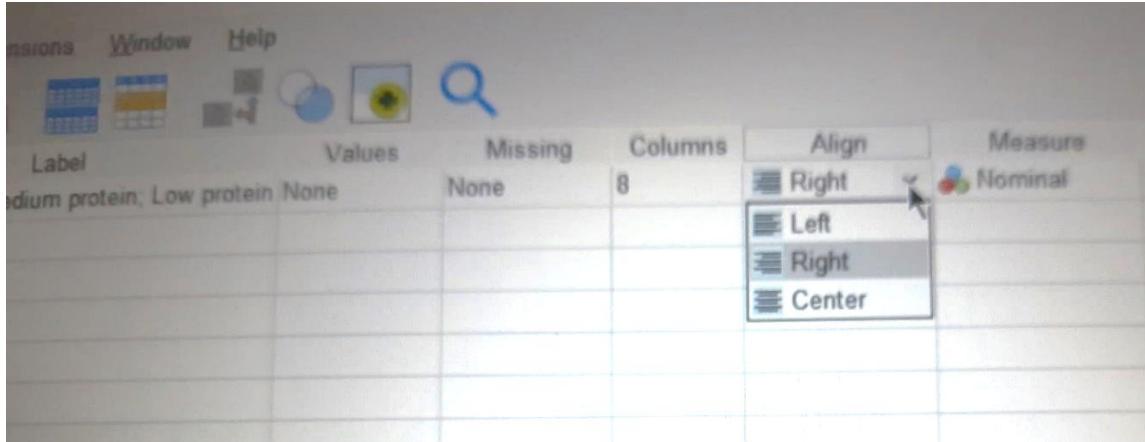
Figure 9.15: Align tab editing in Variable View (SPSS).

9.1.1.1.10 Measure

There are four types of measurement scales: nominal, ordinal, interval , and ratio . While the measurement scale will determine which statistical technique is appropriate for a given set of data, SPSS generally does not discriminate.

Measure is the column where you define the level at which a variable was measured (**Nominal, Ordinal** or **Scale**) (Figure 9.16); SPSS does not distinguish between interval and ratio scales.
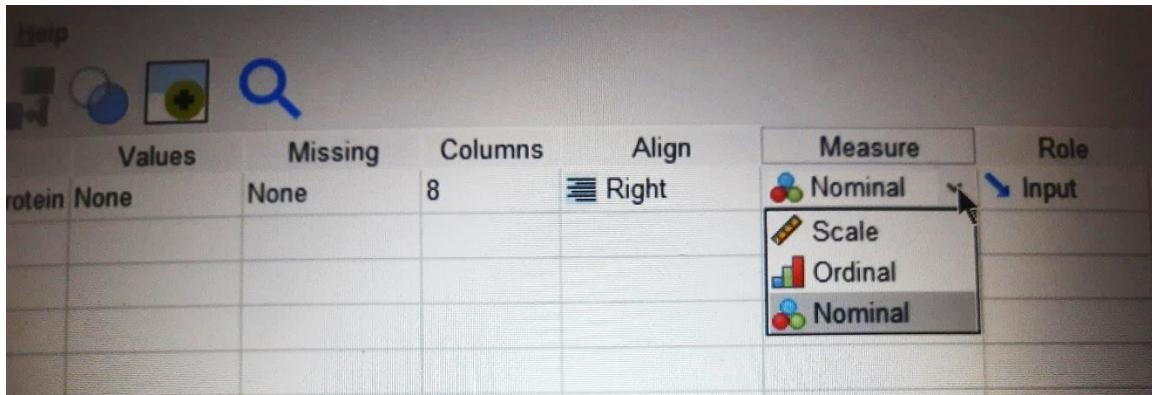
Figure 9.16: Measure tab editing in Variable View (SPSS).

9.1.1.2   Data View

Below is a brief reference guide to each of the menu and some of the options that they **Data View** (Figure 9.17). This is merely a summary.

All the knowledge belongs to God.
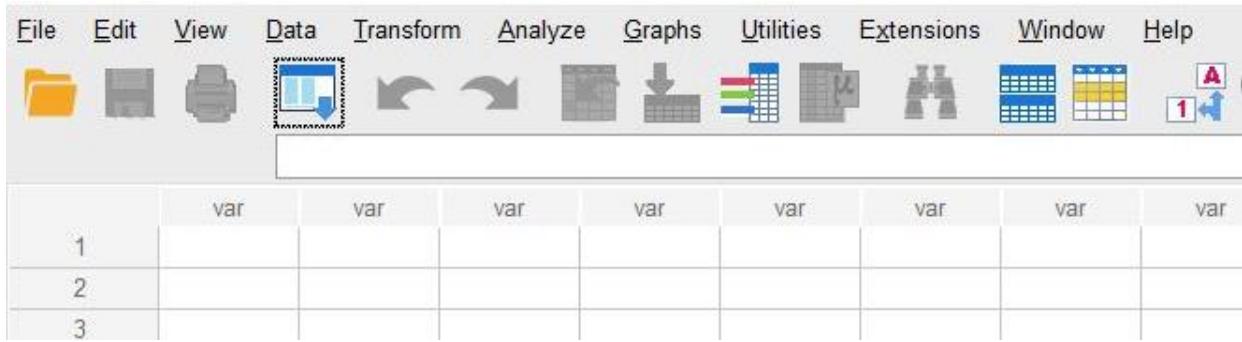
Figure 9.17: Data View menu tabs in SPSS.

9.1.1.2.1 File

This menu allows you to do general things such as saving data, graphs or output. Likewise, you can open previously saved files and print graphs data or output. In essence, it contains all of the options that are customarily found in **File** menus (Figure 9.18).
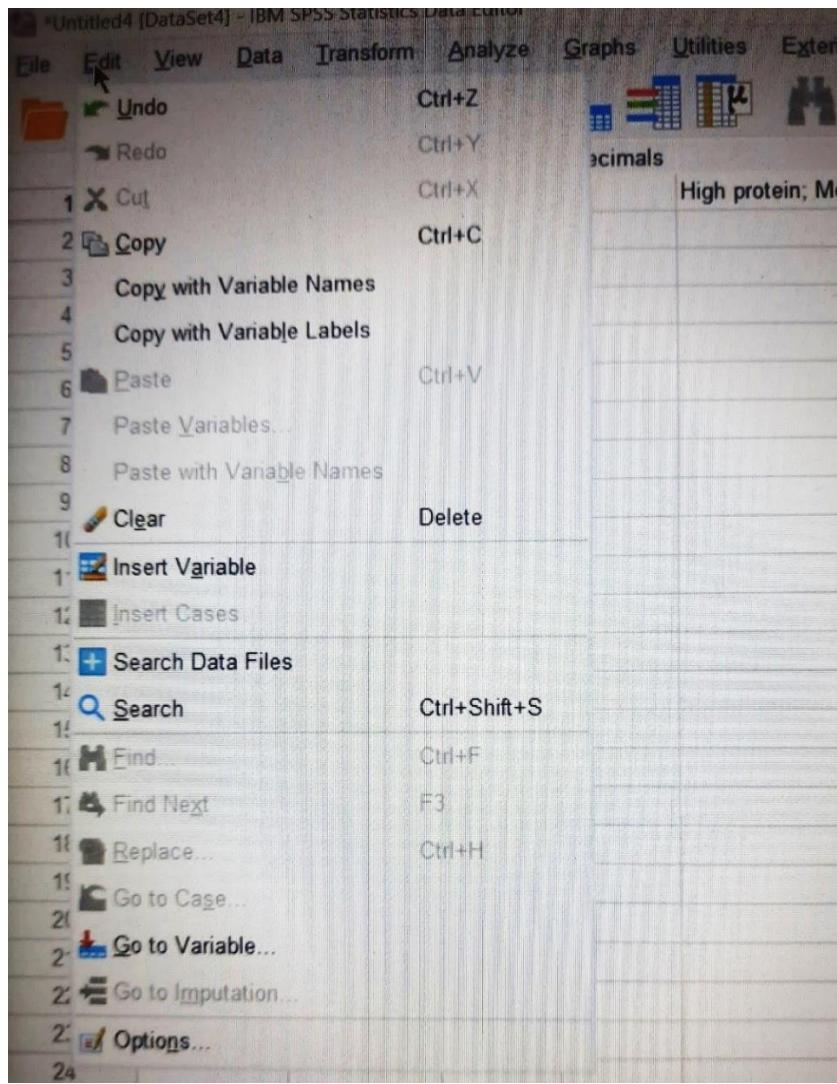
Figure 9.18: File menu in SPSS.

All the knowledge belongs to God.

9.1.1.2.2 Edit

The **Edit** menu contains edit functions for the **Data Editor** (Figure 9.19). In SPSS it is possible to cut and paste blocks of numbers from one part of the data editor to another (which can be very handy when you realize that you've entered lots of numbers in the wrong place). You can also use the **Options…** to select various preferences such as the font that is used for the output. The default preferences are fine for most purposes.
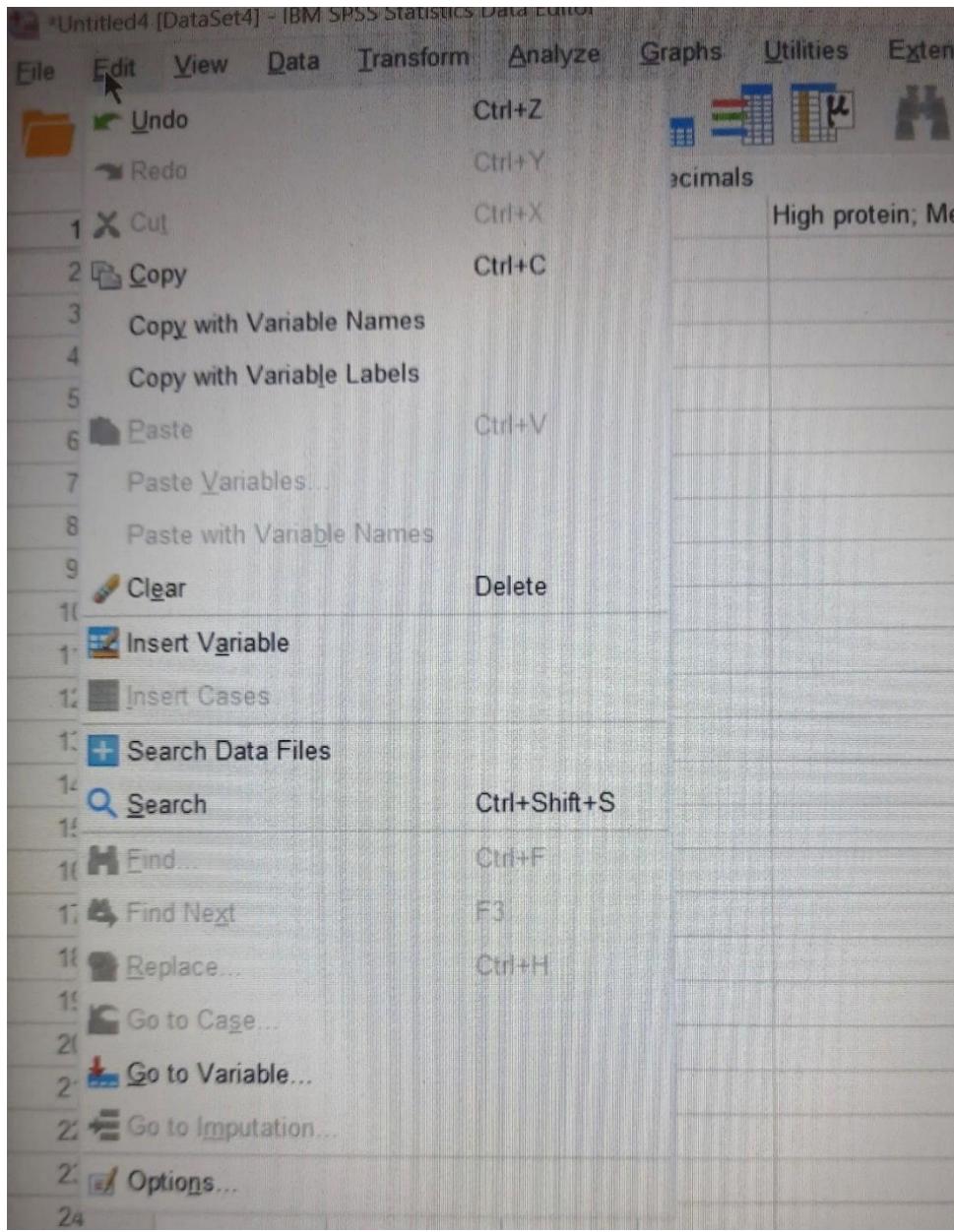
Figure 9.19: Edit menu in SPSS.

9.1.1.2.3 View

This menu (Figure 9.20) deals with system specifications such as whether you have grid lines on the data editor, or whether you display value labels (exactly what value labels are will become clear later).

All the knowledge belongs to God.
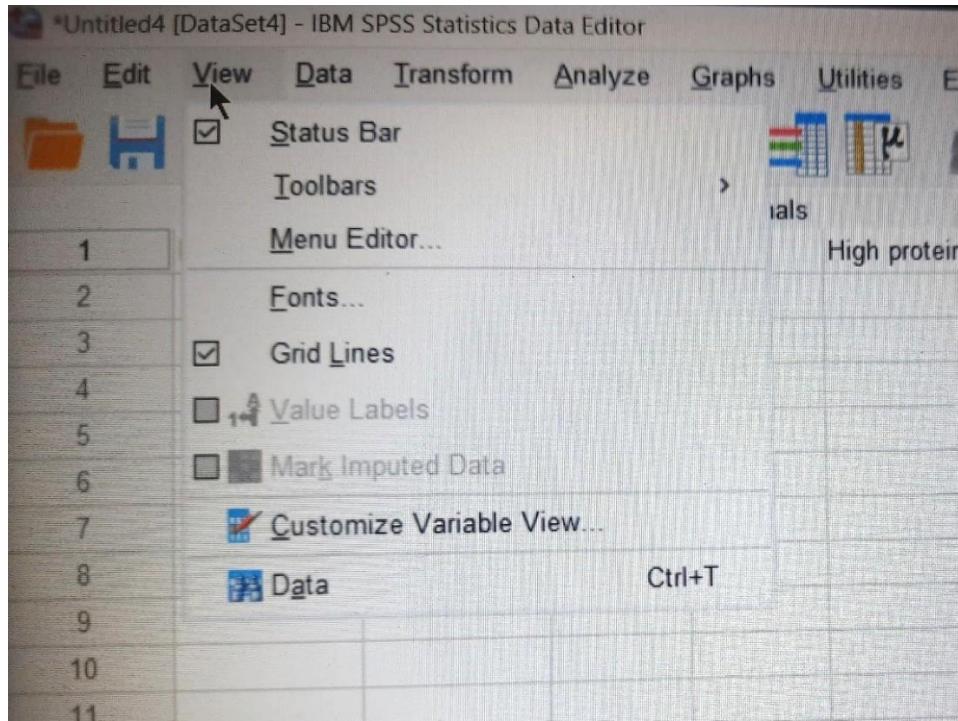
Figure 9.20: View menu in SPSS.

9.1.1.2.4 Data

This menu (Figure 9.21) allows you to make changes to the data editor. We often have more data in a data fi le than we want to include in a specific analysis. For instance, our sample data fi le contains data from four participants, two of whom received special training and two of whom did not. If we wanted to conduct an analysis using only the two participants who did not receive the training, we would need to specify the appropriate subset.
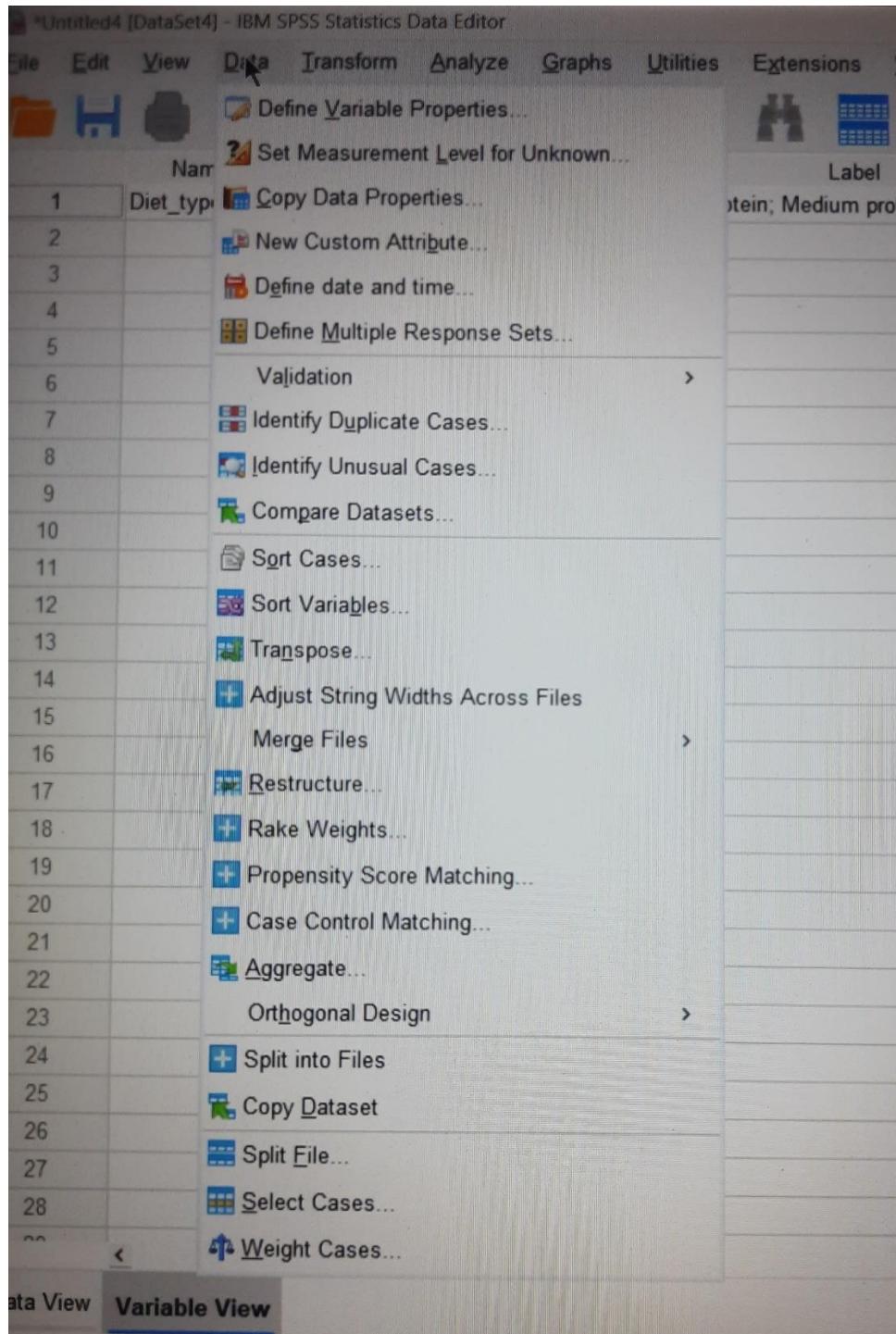
Figure 9.21: Data menu in SPSS.

We can use the **Select Cases** command (at the bottom in above figure) to specify a subset of our data. The **Select Cases** command is located under the **Data** menu. When you select this command, the dialog box below will appear. You can specify which cases (participants) you want to select by using the selection criteria, which appear on the right side of the **Select Cases** dialog box. By default, **All cases** will be selected (Figure 9.22). The most common way to select a subset is to click **If condition is satisfied**, then click on

All the knowledge belongs to God.

the button labeled **If**. This will bring up a new dialog box that allows you to indicate which cases you would like to use.
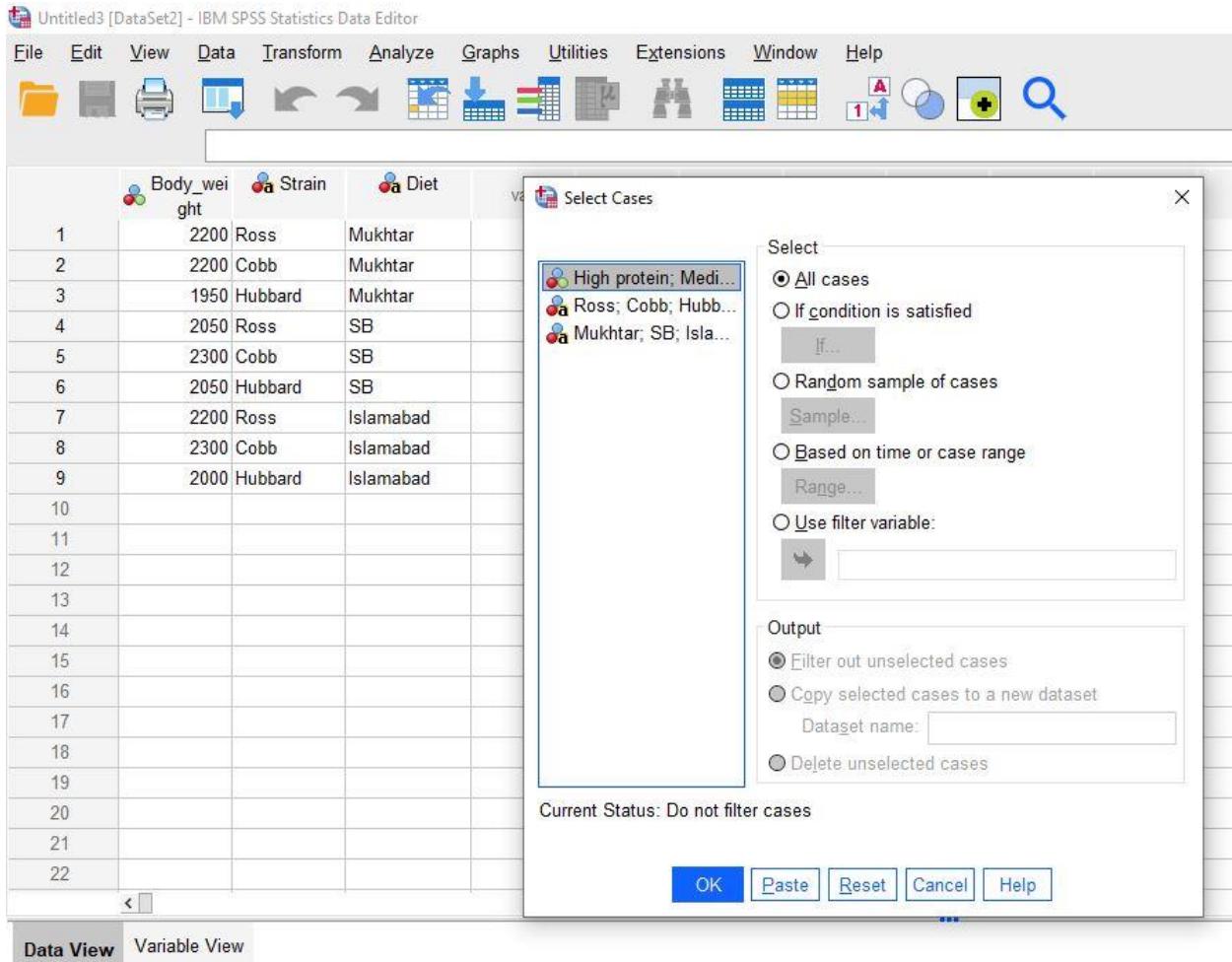
Figure 9.22: Select Case option in Data menu (SPSS).

You can enter the logic used to select the subset in the upper section. If the logical statement is true for a given case, then that case will be selected. If the logical statement is false, that case will not be selected. For instance, you can select all cases that were coded as Body_weight by entering the formula Body_weight=2200 in the upper-left part of the window (Figure 9.23). If weight is 2200, then the statement will be true, and SPSS will select the case. If weight is anything other than 2200, the statement will be false, and the case will not be selected. Once you have entered the logical statement, click **Continue** to return to the **Select Cases** dialog box. Then, click **OK** to return to the data window.

All the knowledge belongs to God.
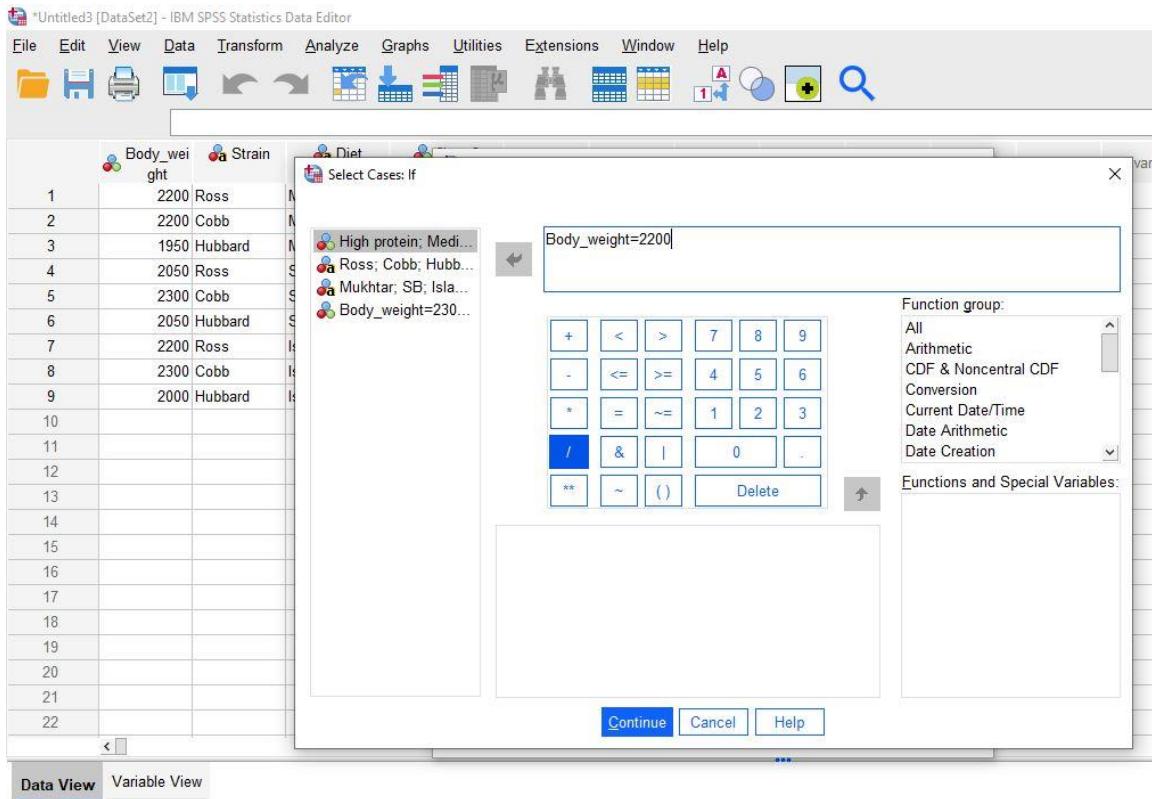
Figure 9.23: Case selection procedure in Data menu (SPSS).

After you have selected the cases, the data window will slightly change. The cases that were not selected will be marked with a diagonal line through the case number. For instance, for our sample data, the first, second and seventh cases are selected (Figure 9.24). All other cases are not selected for this subset:

Figure 9.24: Selected data output using Selected Cases command (SPSS).

All the knowledge belongs to God.

9.1.1.2.5 Transform

You should use this menu if you want to manipulate one of your variables in some way. For example, you can use **Recode into Different Variables** to change the values of certain variables (e.g., if you wanted to adopt a slightly different coding scheme for some reason) (Figure 9.25). The **Compute Variable** function is also useful for transforming data (e.g., you can create a new variable that is the average of two existing variables or addition or subtraction etc.). This function allows you to carry out any number of calculations on your variables.
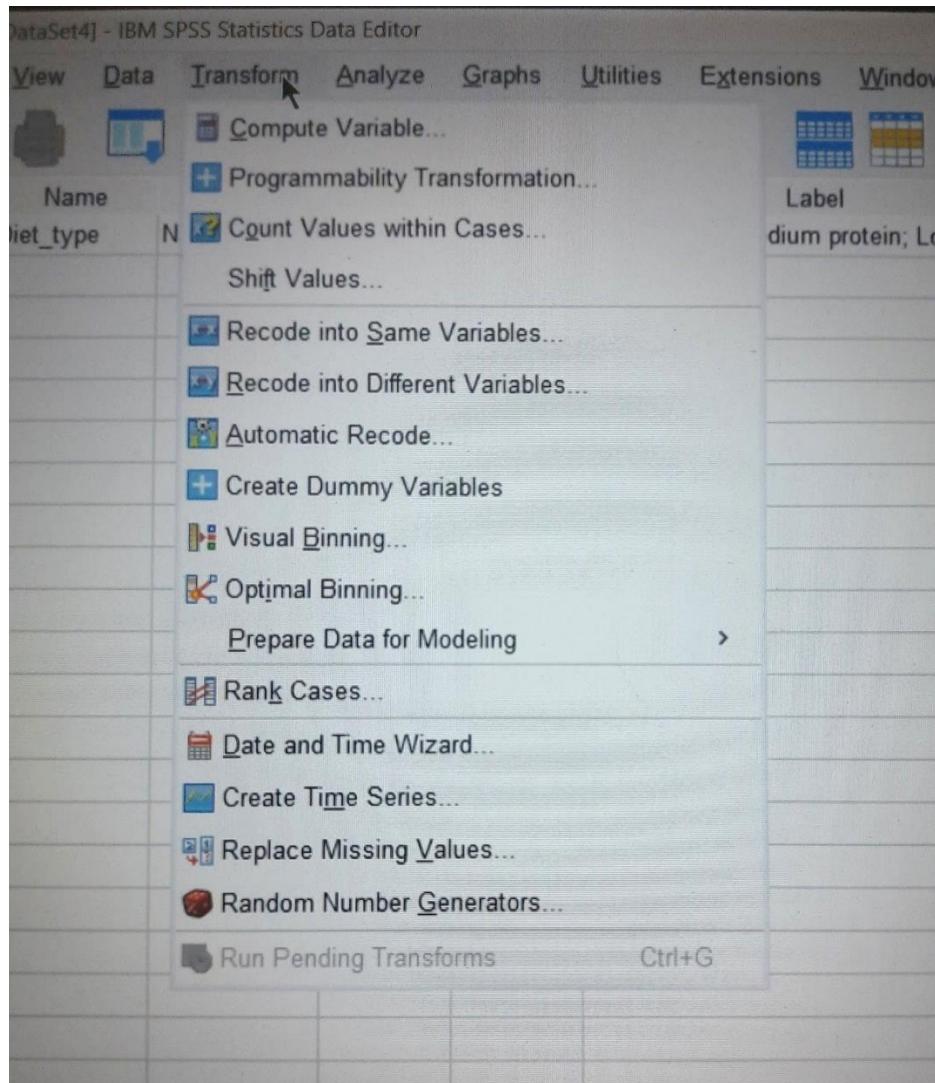
Figure 9.25: Transform menu in SPSS.

SPSS can create a new variable based upon data from another variable. Say we want to split our participants on the basis of their response. We want to create a variable called 'Group', which is coded 1 if the response is low or 2 if the response is high (Figure 9.26). To do this, we click **Transform**, then **Recode into Different Variables**. This will bring up the Recode into Different Variables dialog box shown below.

All the knowledge belongs to God.
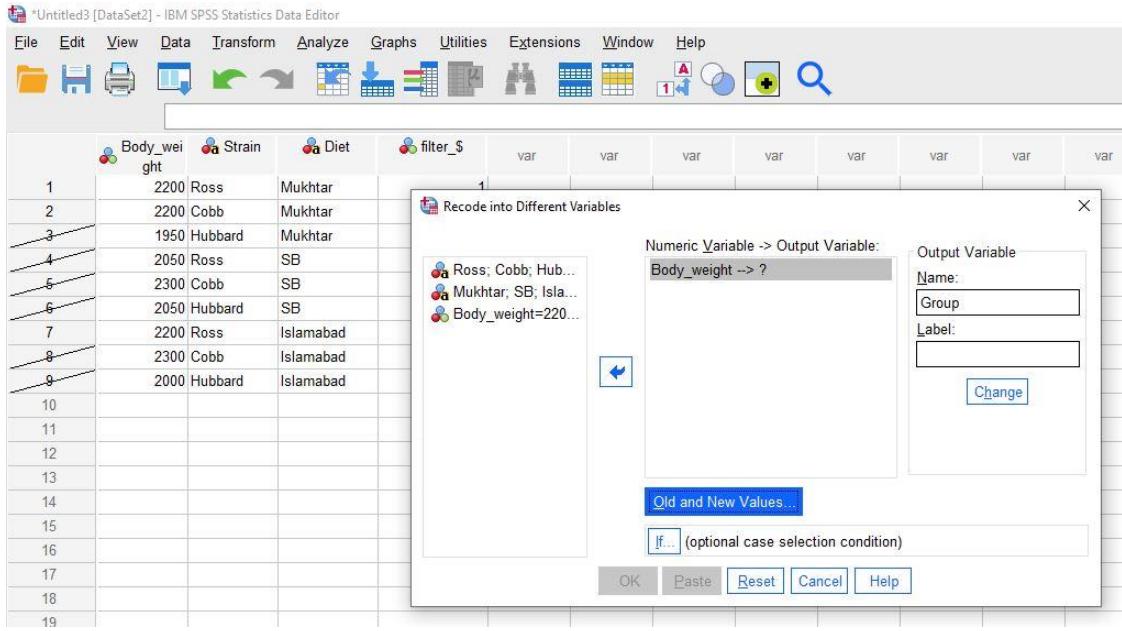
Figure 9.26: Recode function under Transform menu (SPSS).

Click **Old and New Values**. This will bring up the **Recode** dialog box (Figure 9.27).
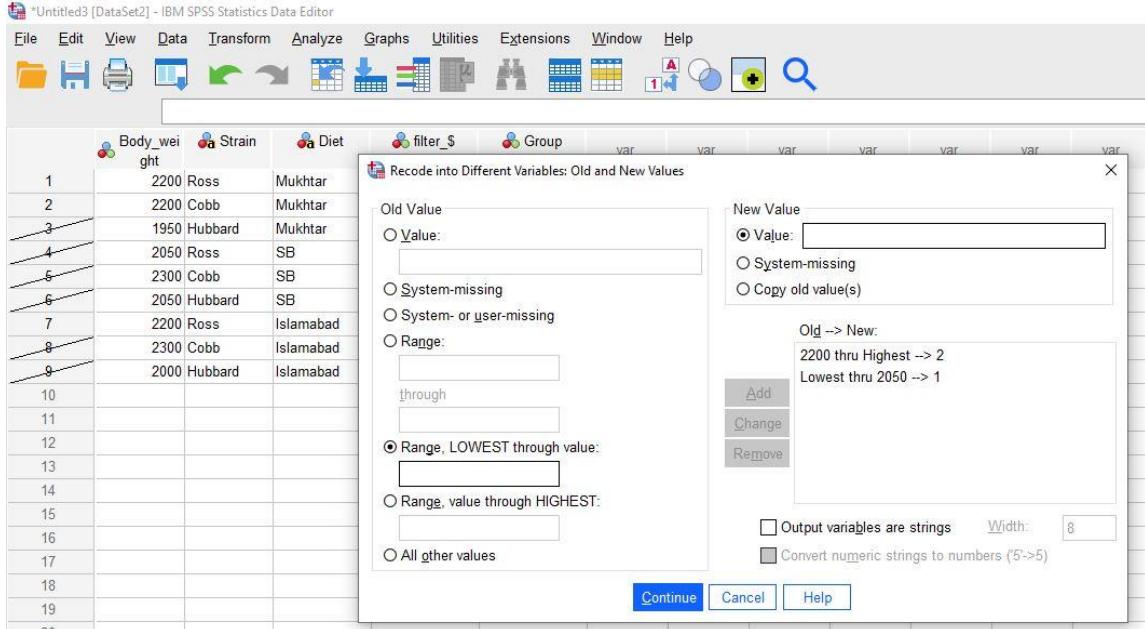
Figure 9.27: Recoding procedure (SPSS).

In the example shown here, we have entered a 2200 in the **Range, value through HIGHEST** field, and a 2 in the value field under **New Value**. When we click **Add**, the blank on the right displays the recoding formula. We next entered 2050 on the left in the **Range, LOWEST through value** blank, and a 1 in the value field under **New Value** (Figure 9.28). Click **Add**, then **Continue**. Click **OK**. You will be redirected to the data

All the knowledge belongs to God.

window shown below. A new variable (Group) will have been added and coded as 1 or 2, based on Body_weight data.

	Body_wei ght	Strain	Diet	filter_$	Group	var	var	var
1	2200	Ross	Mukhtar	1	2.00			
2	2200	Cobb	Mukhtar	1	2.00			
3	1950	Hubbard	Mukhtar	0	1.00			
4	2050	Ross	SB	0	1.00			
5	2300	Cobb	SB	0	2.00			
6	2050	Hubbard	SB	0	1.00			
7	2200	Ross	Islamabad	1	2.00			
8	2300	Cobb	Islamabad	0	2.00			
9	2000	Hubbard	Islamabad	0	1.00			
10								

Figure 9.28: Recoding output (SPSS).

9.1.1.2.6 Analyze

The fun begins here because the statistical procedures lurk in this menu. Below is a brief guide to the options in the statistics menu (Figure 9.29) that will be used.

All the knowledge belongs to God.
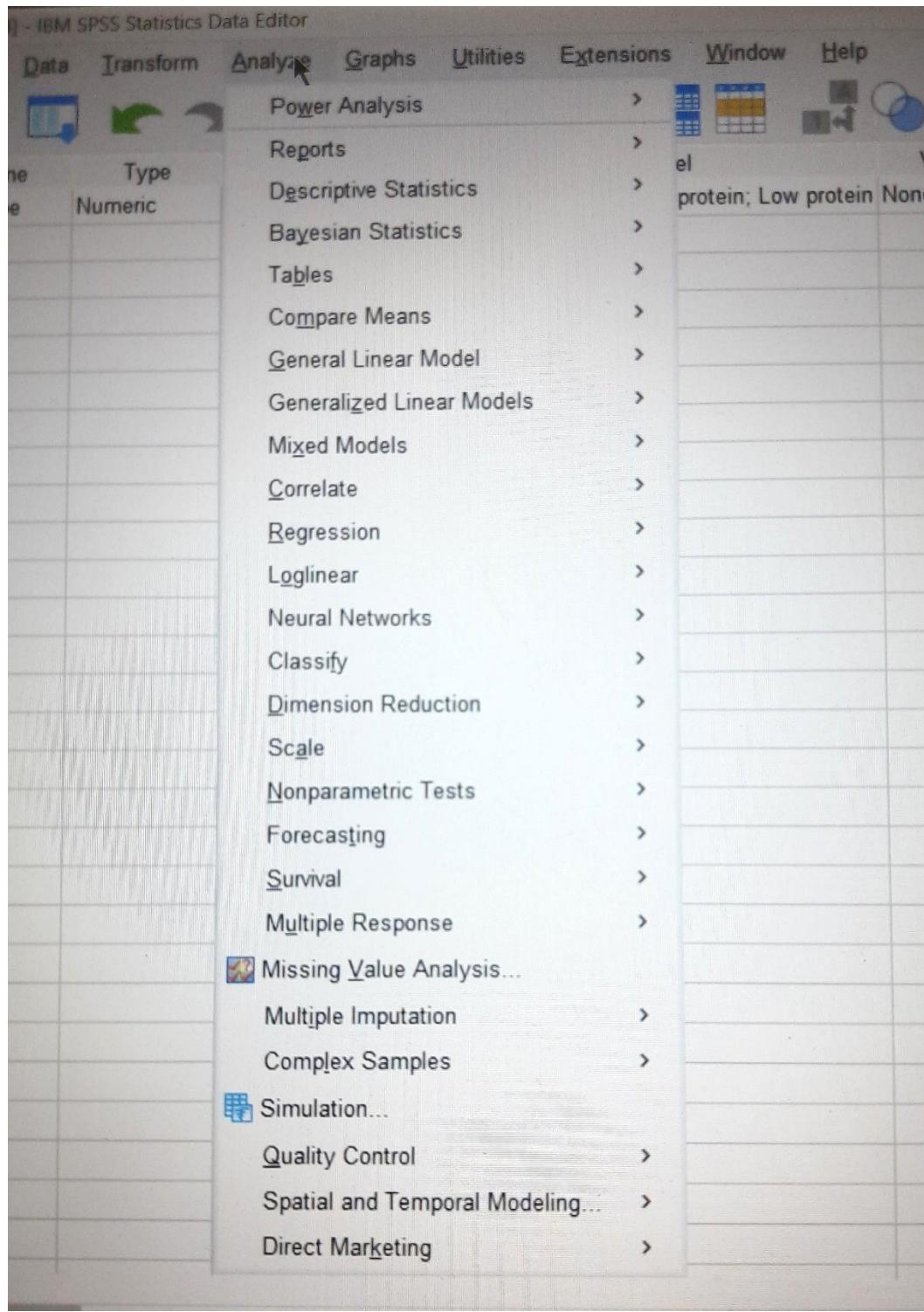
Figure 9.29: Analyze menu in SPSS.

The important features used mostly in research are as following:

Descriptive Statistics: *The procedures used to describe and summarize data are called descriptive statistics*. This menu is for conducting descriptive statistics (mean, mode, median, etc.), frequencies and general data exploration. There is also a command called

crosstabs that is useful for exploring frequency data and performing tests such as chi-square, Fisher's exact test and Cohen's kappa. Moreover, P-P (probability–probability plot) and Q-Q plots are also available under this tab.
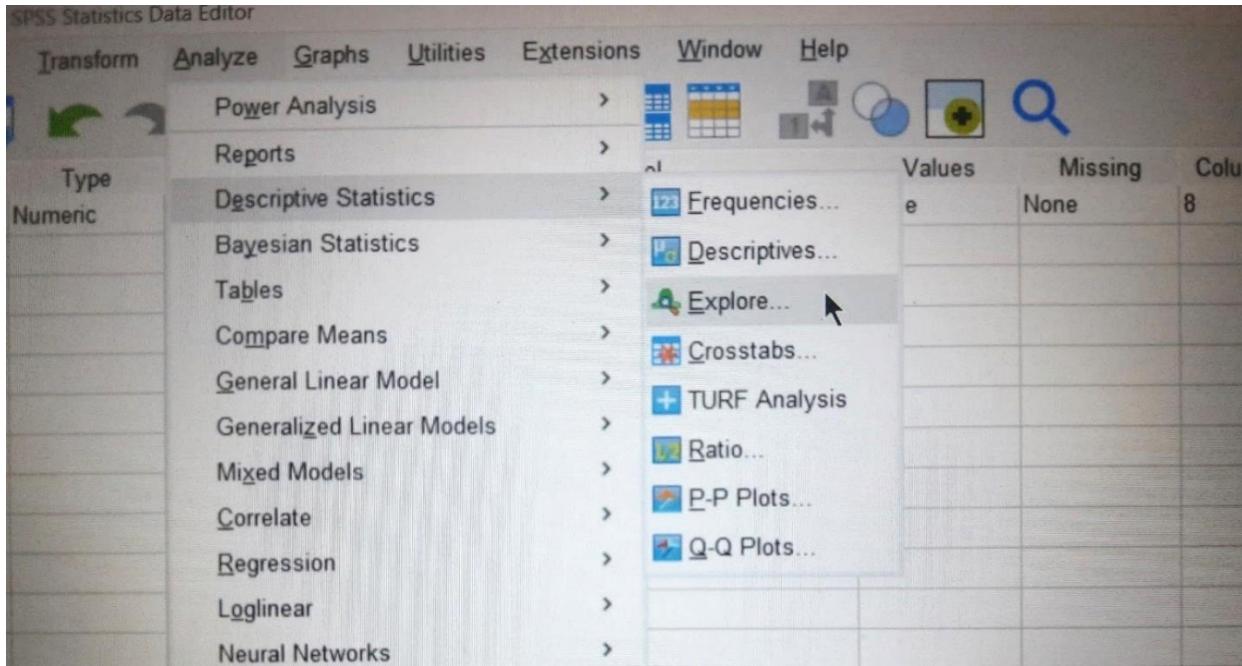
Figure 9.30: Descriptive Statistics option in SPSS.

The **Frequencies** command produces frequency distributions for the specified variables (Figure 9.31). The output includes the number of occurrences, percentages, valid percentages, and cumulative percentages.

The **Frequencies** command will also produce graphical frequency distributions. Click **Frequencies**, and you will be presented with the main dialog box for the **Frequencies** command, where you can enter the variables for which you would like to create graphs or charts using **Charts** option as following:
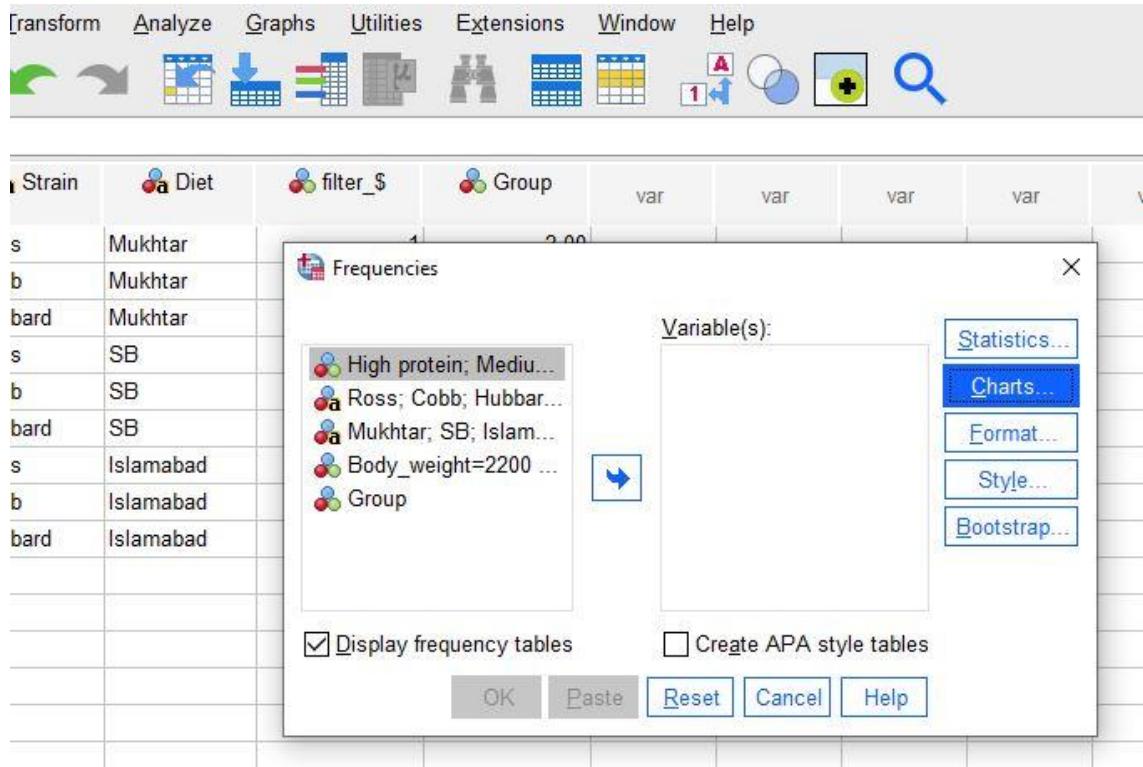
Figure 9.31: Frequency function in Analyze menu (SPSS).

After clicking **Charts**, you will be presented with different options to select from available chart types (Figure 9.32).
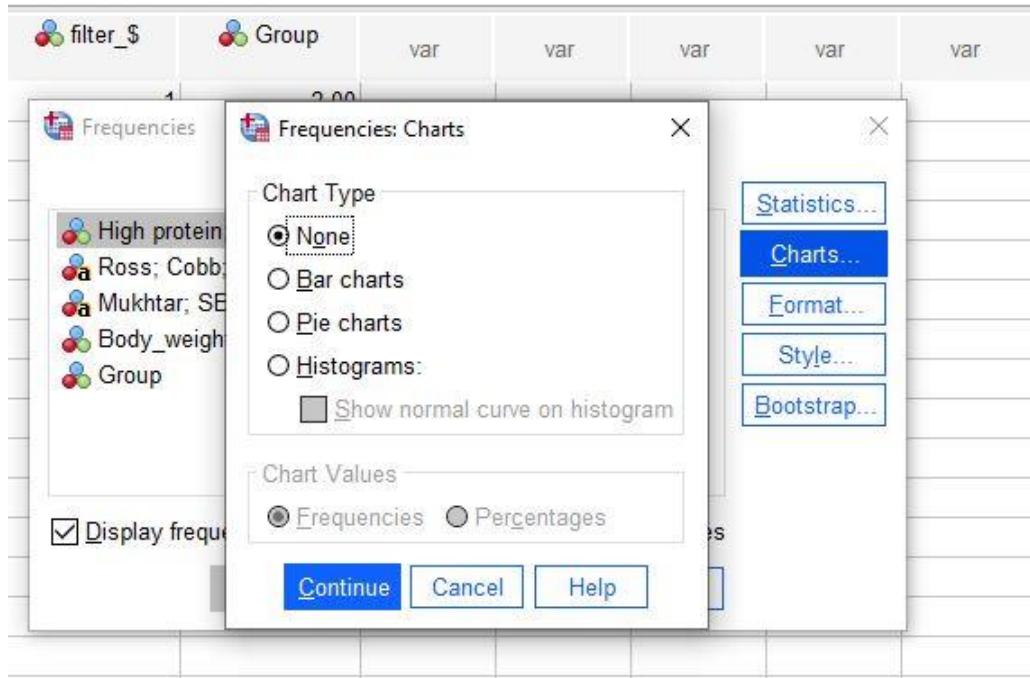
Figure 9.32: Charts selection in Frequency option under Analyze (SPSS).

All the knowledge belongs to God.

Measures of central tendency and measures of dispersion can be calculated using **Descriptives** command.

The **Crosstabs** command produces frequency distributions for multiple variables. The output includes the number of occurrences of each combination of levels of each variable.

Compare Means: This is where you can find t-tests (related and unrelated) and one-way independent ANOVA (Figure 9.33).
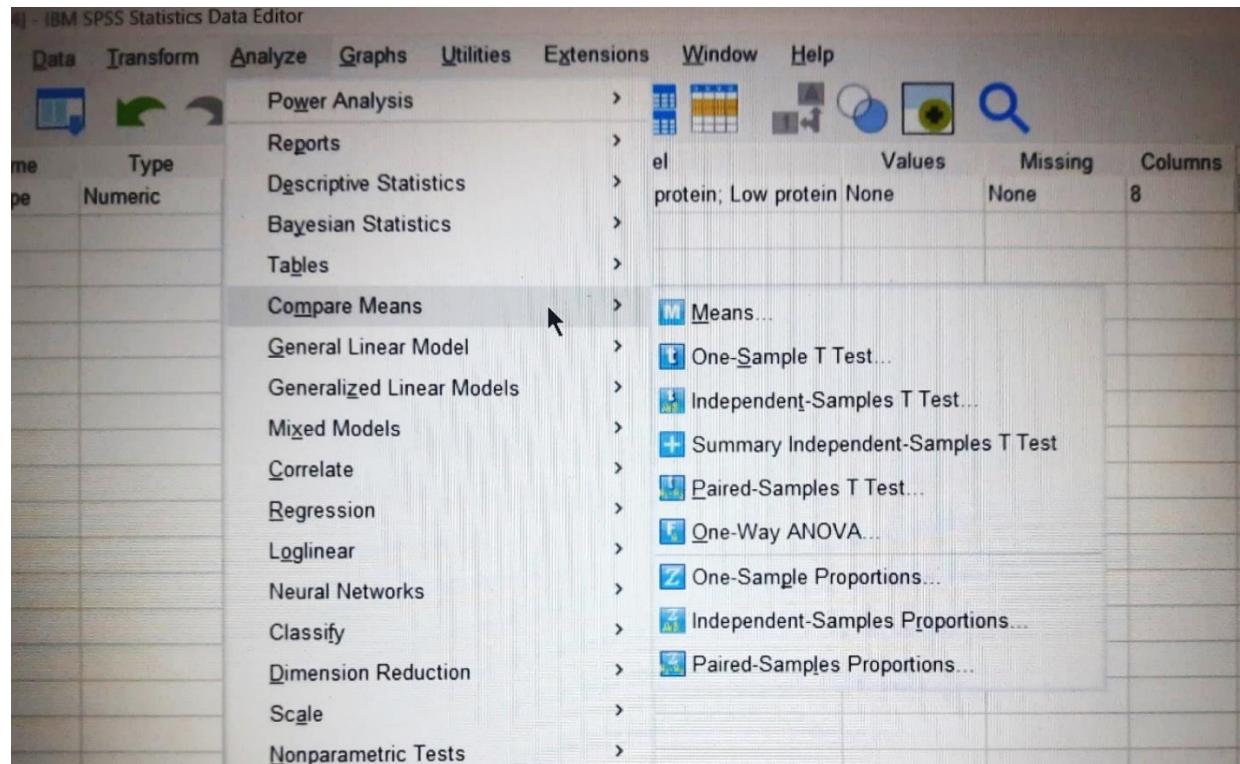
Figure 9.33: Compare Means option in SPSS.

General Linear Model: This menu (Figure 9.34) is for complex ANOVA such as two-way (unrelated, related or mixed), one-way ANOVA with repeated measures and multivariate analysis of variance (MANOVA).

All the knowledge belongs to God.
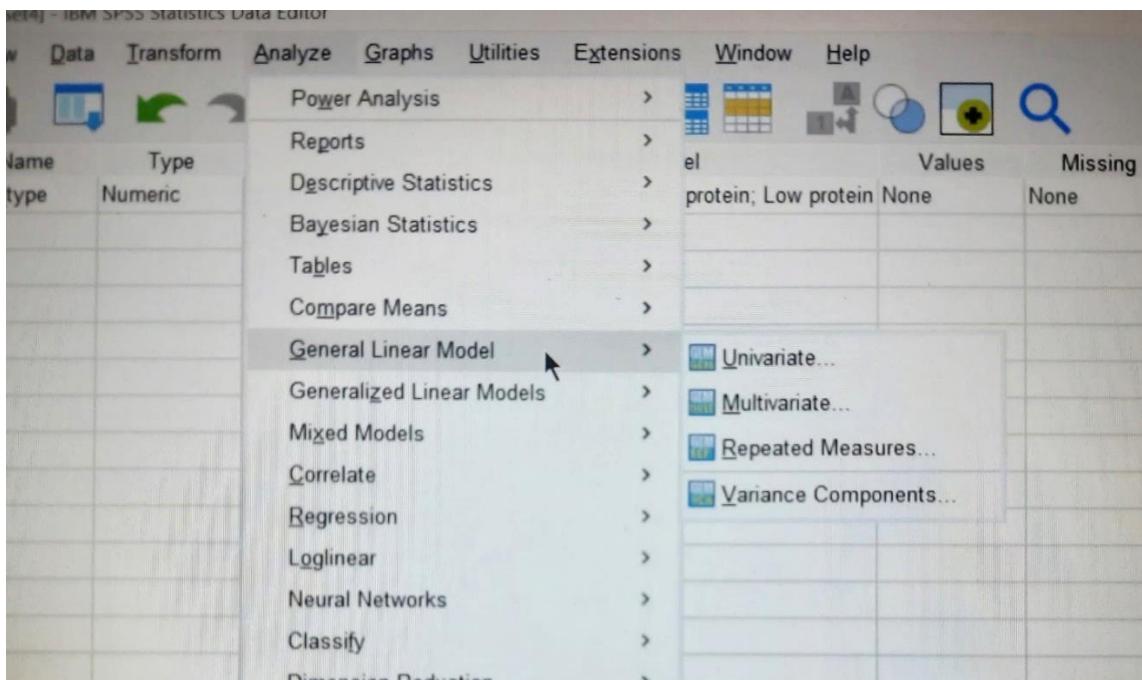
Figure 9.34: General Linear Model (GLM) option in SPSS.

Mixed Models: This menu can be used for running multilevel linear models (MLMs).

Correlate: This is where the correlation techniques are kept! You can do bivariate correlations such as Pearson's R, Spearman's rho (.) and Kendall's tau (t) as well as partial correlations.

The **Pearson correlation coefficient** (sometimes called the Pearson product-moment correlation coefficient or simply the **Pearson r**) determines the strength of the linear relationship between two variables. The correlation coefficient will be between –1.0 and +1.0. Coefficients close to 0.0 represent a weak relationship. Coefficients close to 1.0 or –1.0 represent a strong relationship.

Regression: There are a variety of regression techniques available in SPSS. You can do simple linear regression, multiple linear regression and more advanced techniques such as logistic regression.

Linear… allows for the prediction of one variable from another.

R Square (called the coefficient of determination) in output of linear regression gives you the proportion of the variance of your dependent variable that can be explained by variation in your independent variable.

The multiple linear regression analysis is also possible using **Linear** command. It allows the prediction of one variable from several other variables.

Loglinear: Loglinear analysis is hiding in this menu, waiting for you, and ready to pounce like a tarantula from its burrow.

Nonparametric Tests: There are a variety of non-parametric statistics available (Figure 9.35) such as the chi-square goodness-of-fit statistic, the binomial test, the Mann–Whitney test, the Kruskal–Wallis test, Wilcoxon's test and Friedman's ANOVA.

All the knowledge belongs to God.
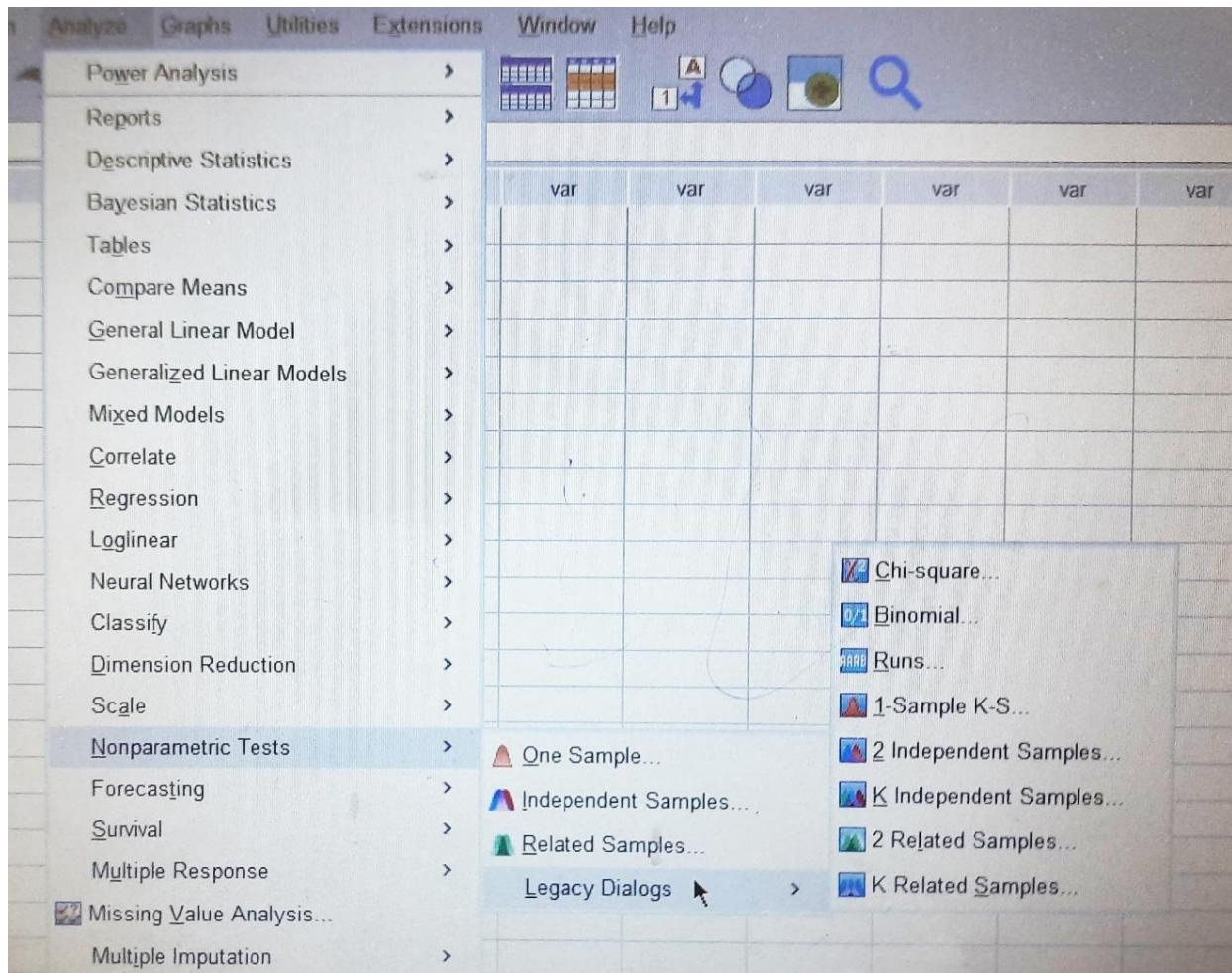
Figure 9.35: Non-parametric Tests option in SPSS.

9.1.1.2.7 Graphs

SPSS has some graphing facilities, and this menu is used to access the submenu **Chart Builder** (Figure 9.36). The types of graphs you can do include bar charts, histograms, scatterplots, box–whisker plots, pie charts and error bar graphs to name but a few.
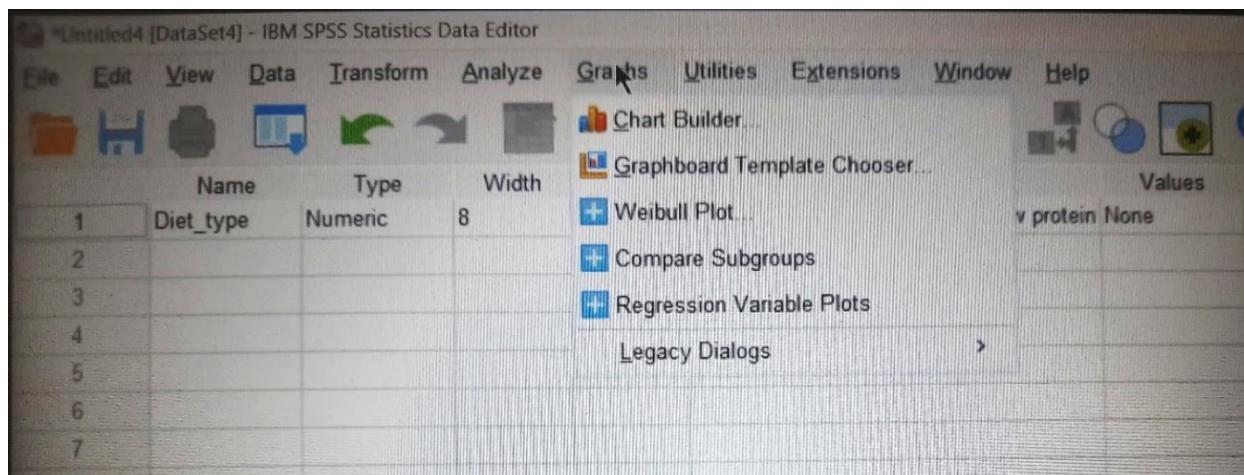
Figure 9.36: Graphs menu in Data View (SPSS).

All the knowledge belongs to God.

The **Chart Builder** command will produce graphical frequency distributions. You will be presented with the main dialog box for the **Chart Builder** command (Figure 9.37), where you can enter the variables for which you would like to create graphs or charts:
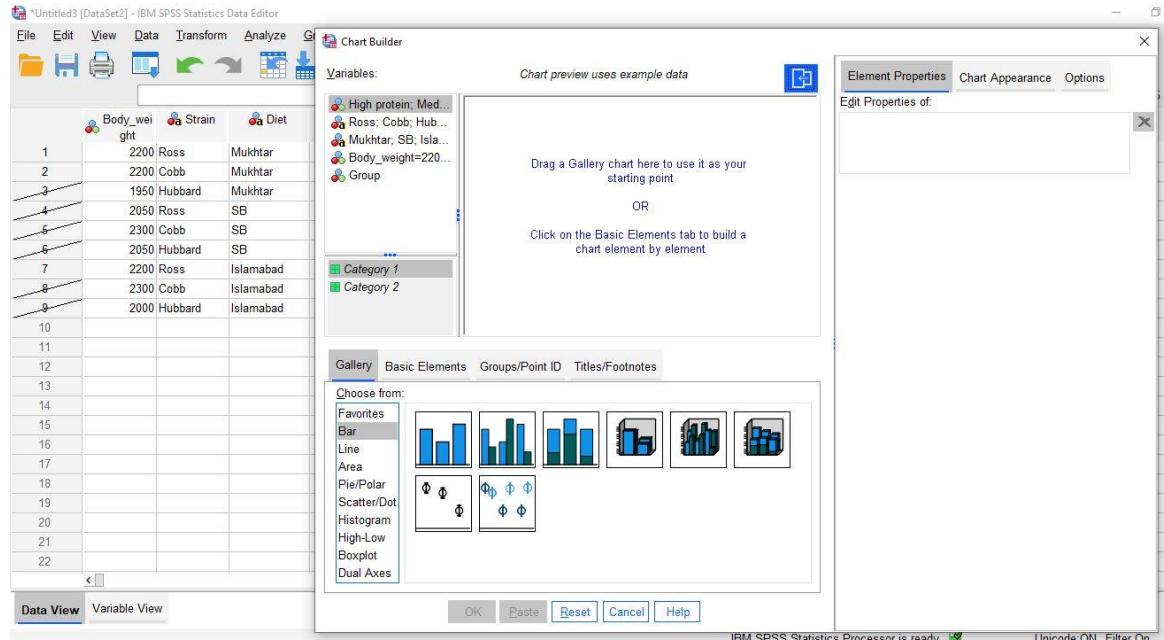
Figure 9.37: Chart builder window in Graphs menu (SPSS).

Near the middle of the dialog box, there are four main tabs that let you control the graphs you are making. The first one is the **Gallery** tab (Figure 9.38). The Gallery tab allows you to choose the basic format of your graph as following:
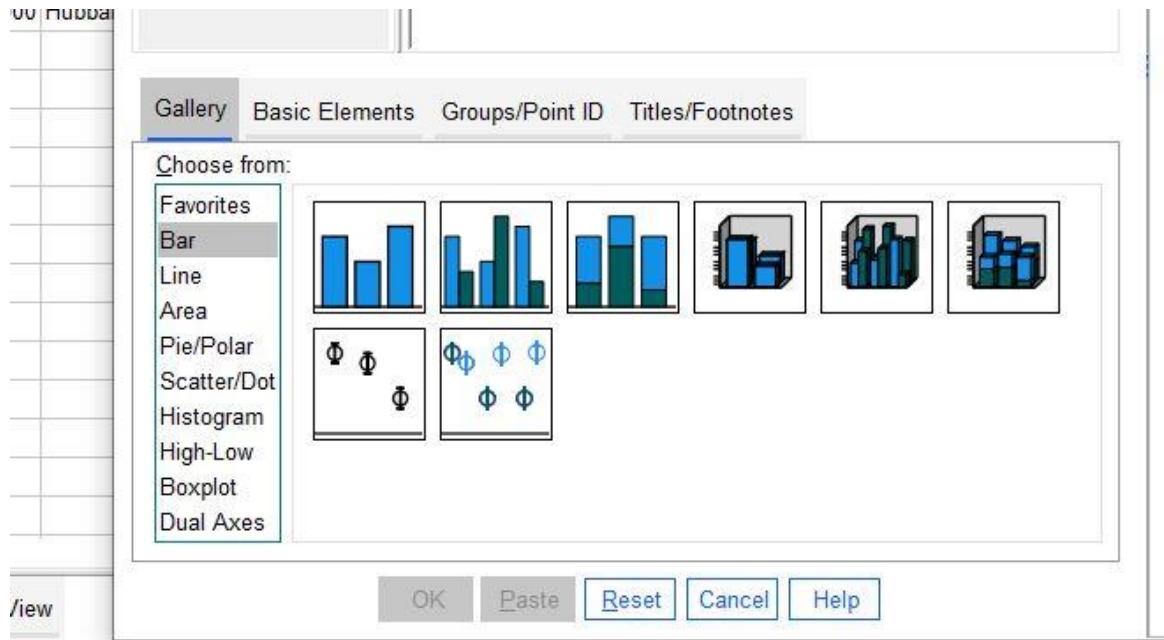
Figure 9.38: Gallery selection in Graphs menu (SPSS).

All the knowledge belongs to God.

Alternatively, you can use the **Basic Elements** tab to drag a coordinate system (labeled **Choose Axes**) to the top window, then drag variables and elements into the window. The other tabs (**Groups/Point ID** and **Titles/Footnotes**) can be used for adding other standard elements to your graphs. Even though the scatterplot is a two-dimensional graph, it can plot a third variable. To make it do so, select the **Groups/Point ID** tab in the **Chart Builder**.

9.1.1.2.8 Utilities

In this menu there is an option, '**Data File Comments**,' that allows you to comment on your data set (Figure 9.39). This can be quite useful because you can write yourself notes about from where the data come, or the date they were collected and so on.
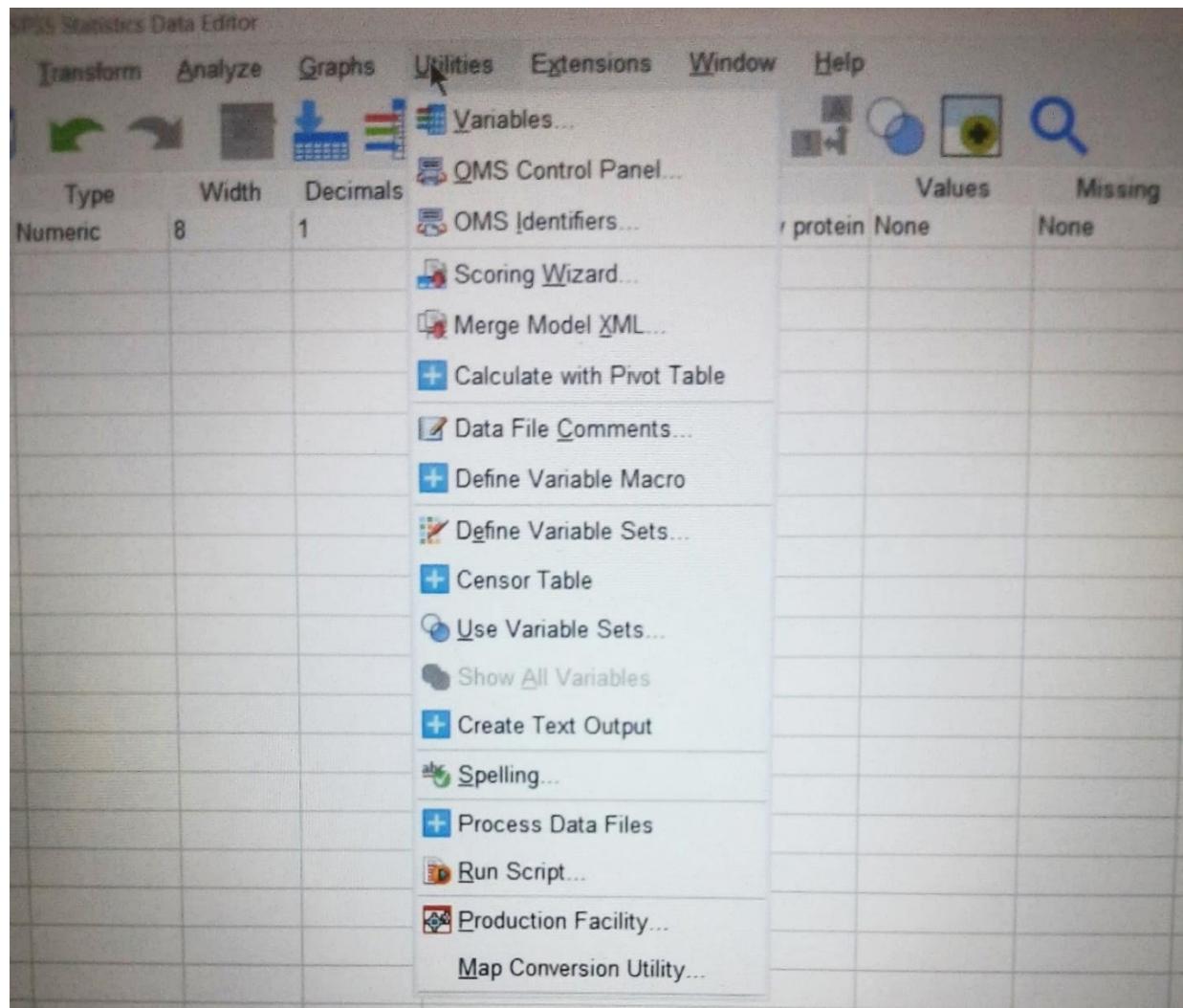
Figure 9.39: Utilities menu in SPSS.

9.1.1.2.9 Extensions or Add-on

SPSS sells several add-ons that can be accessed through this menu (Figure 9.40). For example, SPSS has a program called Sample Power that computes the sample size required for studies, and power statistics.

All the knowledge belongs to God.
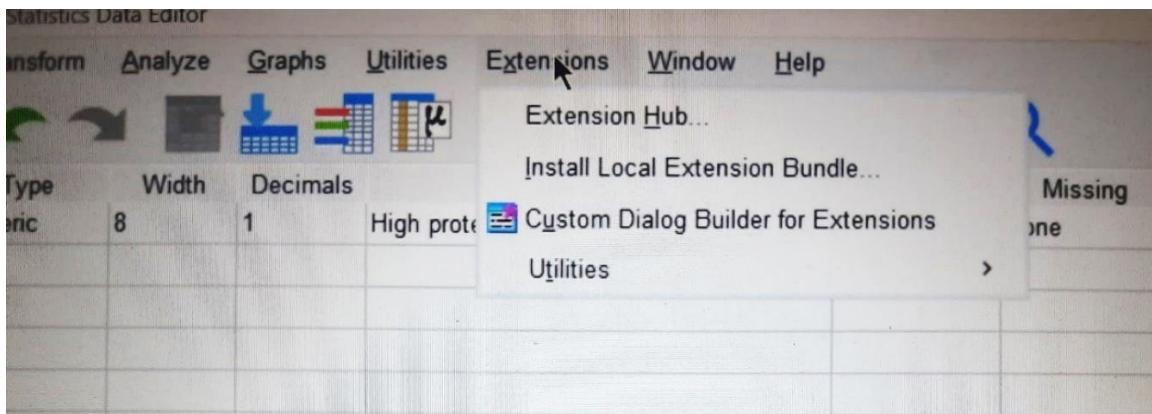
Figure 9.40: Extension menu in SPSS.

9.1.1.2.10 Window

This menu allows you to switch from window to window (Figure 9.41). So, if you're looking at the output and you wish to switch back to your data sheet, you can do so using this menu. There are icons to shortcut most of the options in this menu, so it isn't particularly useful.
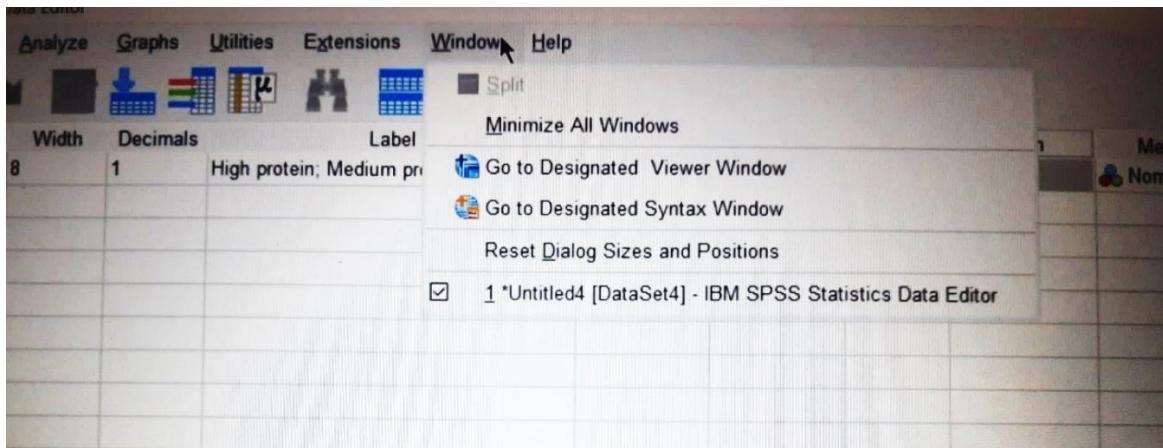
Figure 9.41: Window menu in SPSS.

9.1.1.2.11 Help

This is an invaluable menu because it offers you online help on both the system itself and the statistical tests (Figure 9.42). The statistics help files are incomprehensible at times.

All the knowledge belongs to God.
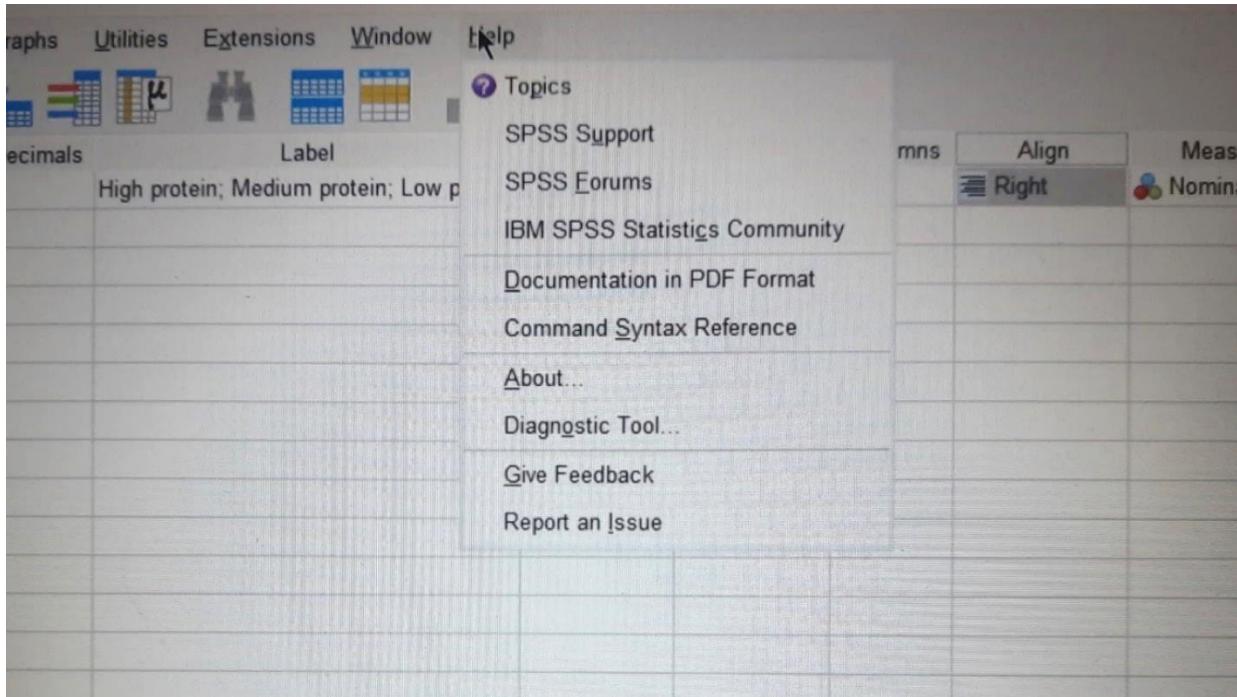
Figure 9.42: Help menu in SPSS.

9.1.2: Entering Data into Data View

Suppose we have data of body weight (Table 9.2) of three chicken strains (Ross, Hubbard and Cobb) due to three diets (Mukhtar; SB and Islamabad):

Table 9.2: Sample question for data entry into Data View (SPSS).

Diet	Ross	Cobb	Hubbard
Mukhtar	2000	2200	1950
SB	2050	2300	2050
Islamabad	2200	2300	2000

9.1.2.1 SPSS Entries

The first variable in our data set is the name of the feed (diet). This variable consists of names; therefore, it is a string variable. Similarly, the strain are string variable. The **Variable View** (Figure 9.43) need to be edited for these entries:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Body_weight	Numeric	8	0	Response	None	None	8	Right	Scale	Input
2	Strain	String	8	0	Ross; Cobb; Hubbard	{1, Ross}...	None	8	Left	Nominal	Input
3	Diet	String	9	0	Mukhtar; SB; Islamabad	{1, Mukhtar}...	None	9	Left	Nominal	Input
4											

Figure 9.43: Sample data entry into Variable View (SPSS).

The **Label** will be used for strains and diet coding as following:

All the knowledge belongs to God.
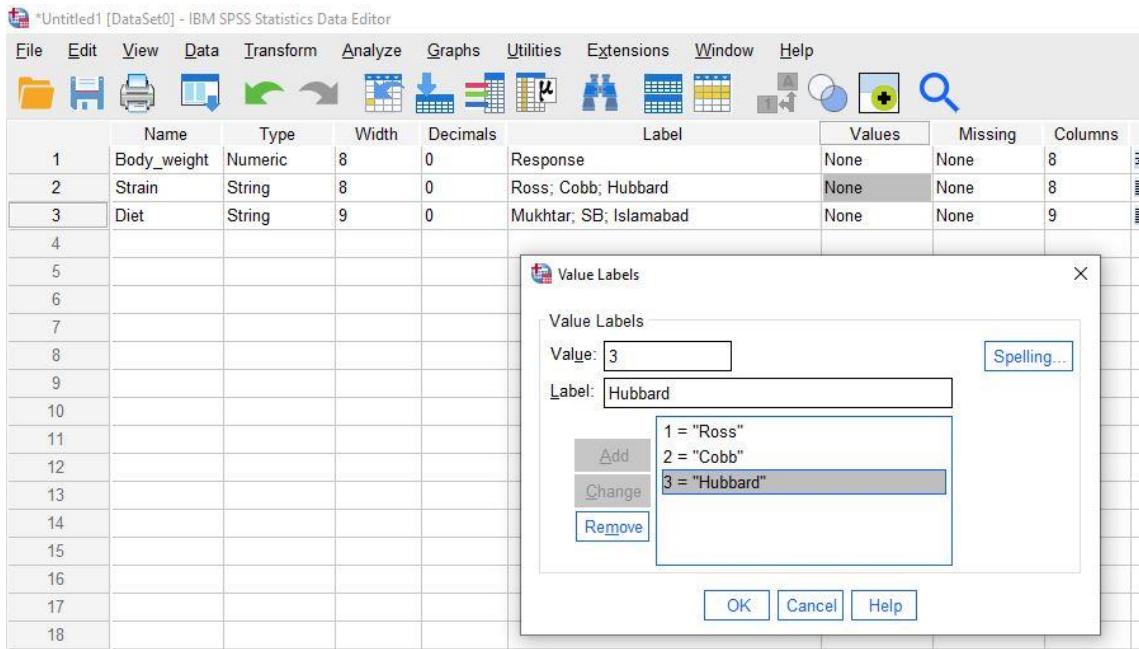
Figure 9.44: Label editing for sample data (SPSS).

The data after entering will look like:

Figure 9.45: Data entry into Data View (SPSS).

9.1.2.2 SPSS Output

After an analysis is performed, the output is placed in the output window, and the output window becomes the active window. If this is the first analysis you have conducted since starting SPSS, then a new output window will be created. If you have run previous analyses and saved them, your output is added to the end of your previous output.

The output window is split into two sections (Figure 9.46). The left section is an outline of the output (SPSS refers to this as the **outline view**). The right section is the output itself.

All the knowledge belongs to God.
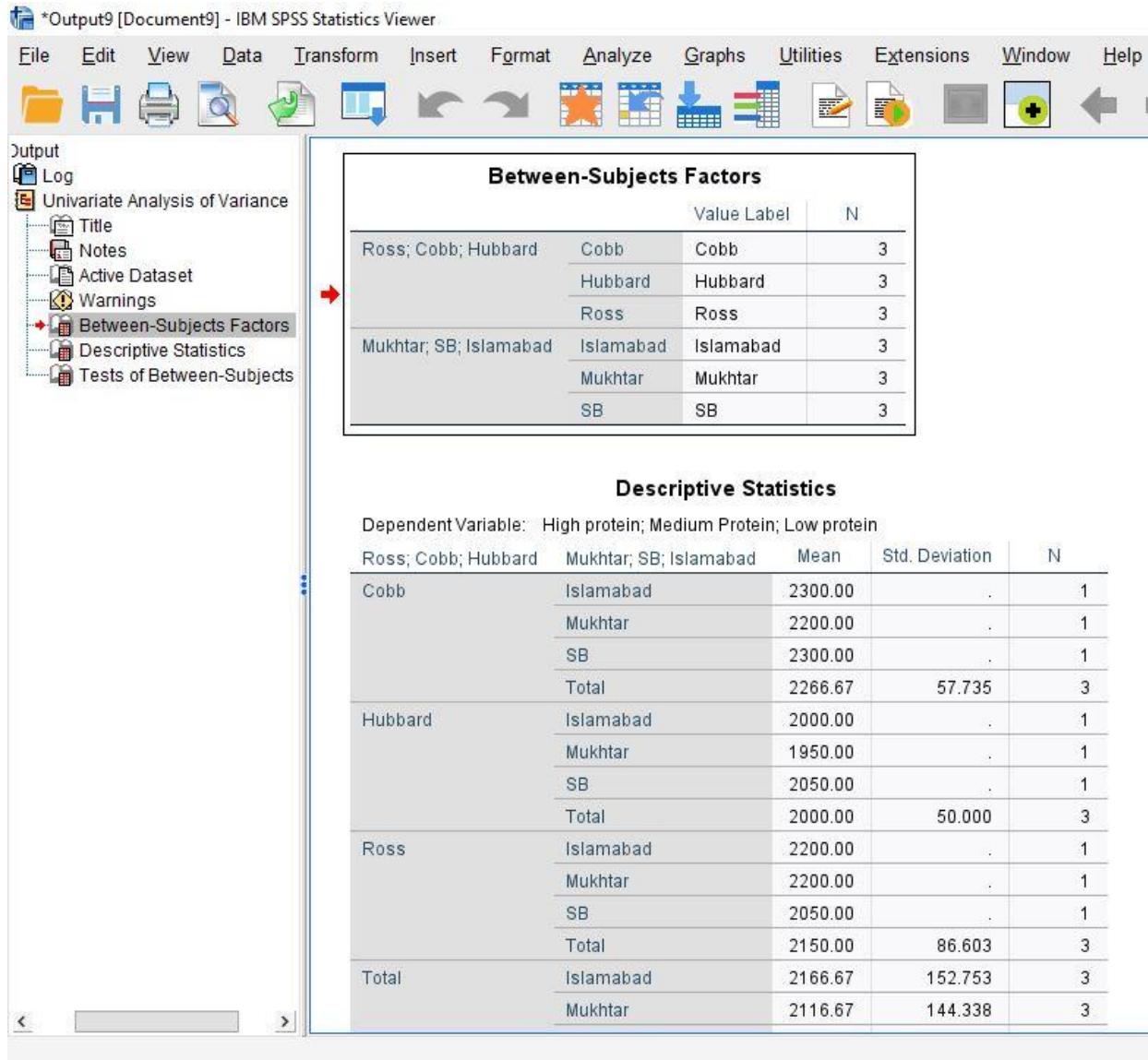
Figure 9.46: SPSS output.

The section on the left of the output window provides an outline of the entire output window. All of the analyses are listed in the order in which they were conducted. Note that this outline can be used to quickly locate a section of the output. Simply click on the section you would like to see, and the right window will jump to the appropriate place.

To print your output, simply click **File**, then **Print**, or click on the printer icon on the toolbar. You will have the option of printing all of your output or just the currently selected section. Be careful when printing! Each time you run a command, the output is added to the end of your previous output. Thus, you could be printing a very large output file containing information you may not want or need.

9.1.2.3   File Saving and Retrieving

The saving file is just like MS Word or any other document. The data and output both can be saved (**waqas.sav** in this case) in this procedure in desired place (Figure 9.47) using following commands:

All the knowledge belongs to God.
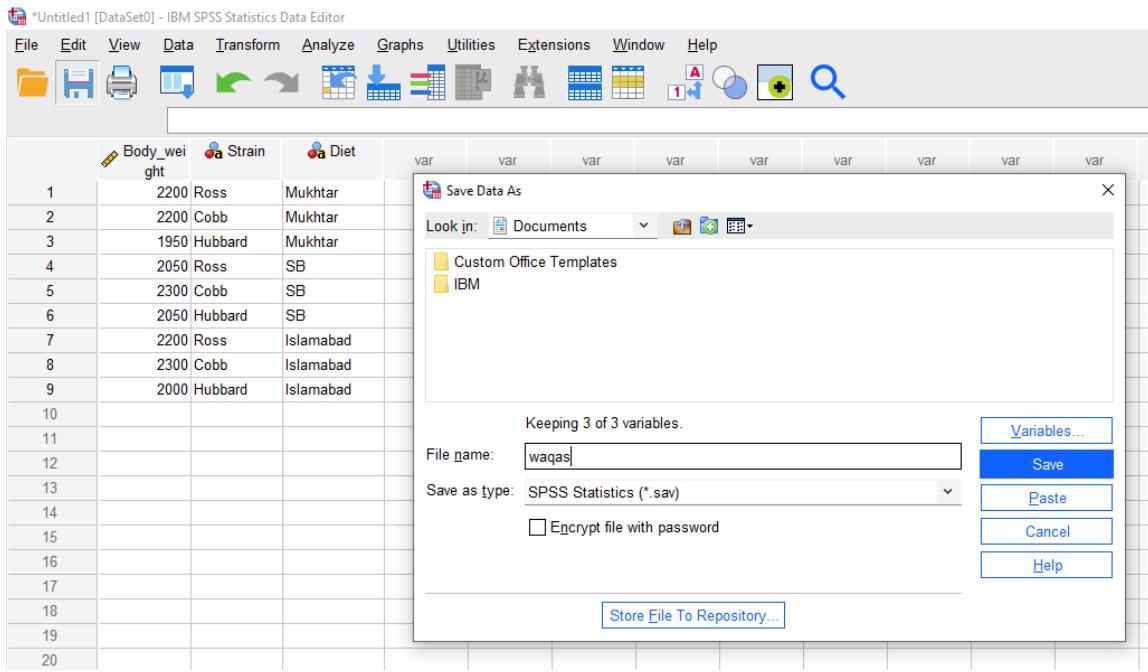
File → Save as

Figure 9.47: File saving in SPSS.

The saved file can be retrieved using the same path by using **File** tab from menu using following commands:

File → Open → Data
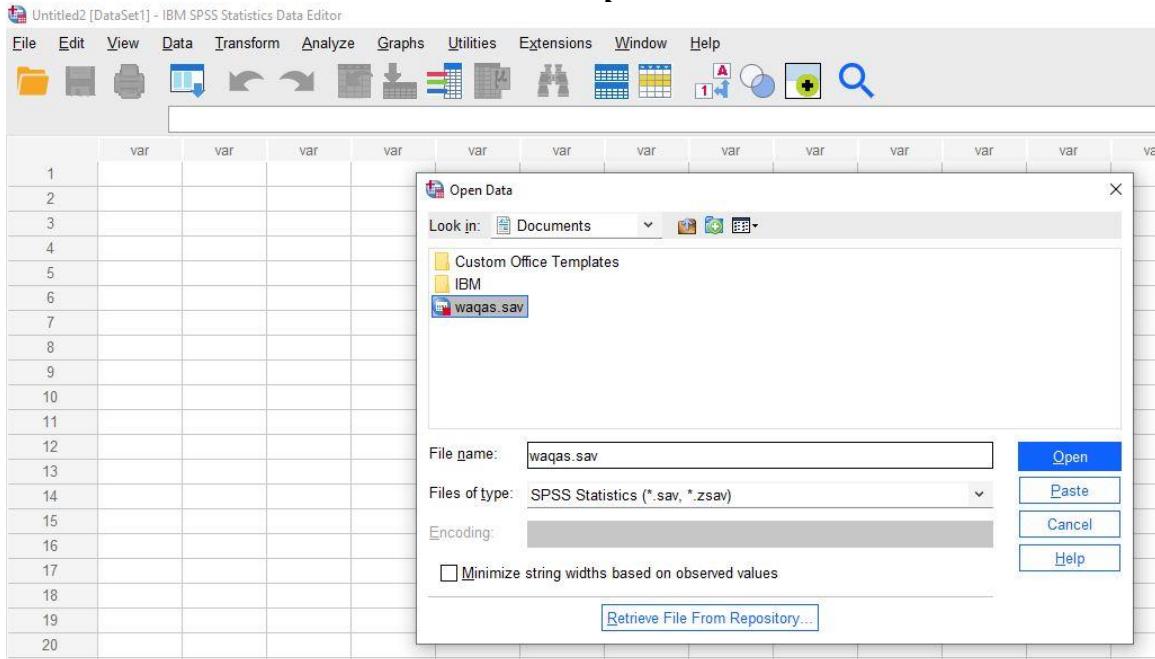
Figure 9.48: File retrieving in SPSS.

Your file (**waqas.sav** in this case) will be retrieved by this procedure.

All the knowledge belongs to God.

9.2: P-VALUE AND ITS SIGNIFICANCE

9.2.1: Probability

The classical probability dates back to the 17th century and the work of two mathematicians, Pascal and Fermat. Much of this theory developed out of attempts to solve problems related to games of chance, such as those involving the rolling of dice. Examples from games of chance illustrate very well the principles involved in classical probability. For example, if a fair six-sided die is rolled, the probability that a 1 will be observed is equal to $1 = 6$ and is the same for the other five faces. If a card is picked at random from a well-shuffled deck of ordinary playing cards, the probability of picking a heart is $13 = 52$.

In the early 1950s, LJ. Savage gave considerable impetus to what is called the **personalistic** or subjective concept of probability. This view holds that probability measures the confidence that a particular individual has in the truth of a particular proposition. This concept does not rely on the repeatability of any process. In fact, by applying this concept of probability, one may evaluate the probability of an event that can only happen once, for example, the probability that a cure for cancer will be discovered within the next 10 years.

9.2.2: P-Value

The p value is a number that tells us how unusual our sample results are, given that the null hypothesis is true. A p value indicating that the sample results are not likely to have occurred, if the null hypothesis is true, provides justification for doubting the truth of the null hypothesis. To judge whether an observed value (t_s) is "far" in the tail of the t distribution, we need a quantitative yardstick for locating t_s within the distribution. This yardstick is provided by the P-value. *The P-value for a hypothesis test is the probability, computed under the condition that the null hypothesis is true, of the test statistic being at least as extreme as the value of the test statistic that was actually obtained.*

Thus, a P-value conveys much information about the weight of evidence against H_o, and so a decision maker can draw a conclusion at any specified level of significance. More formally, we define the P-value as the smallest level of significance that would lead to rejection of the null hypothesis H_o. The P-value, which is sometimes abbreviated as simply "P," is the shaded area in Figure 9.49. Note that we have defined the P-value as the total area in both tails; this is sometimes called the **"two-tailed" P-value**.

Figure 9.49: P value.

All the knowledge belongs to God.
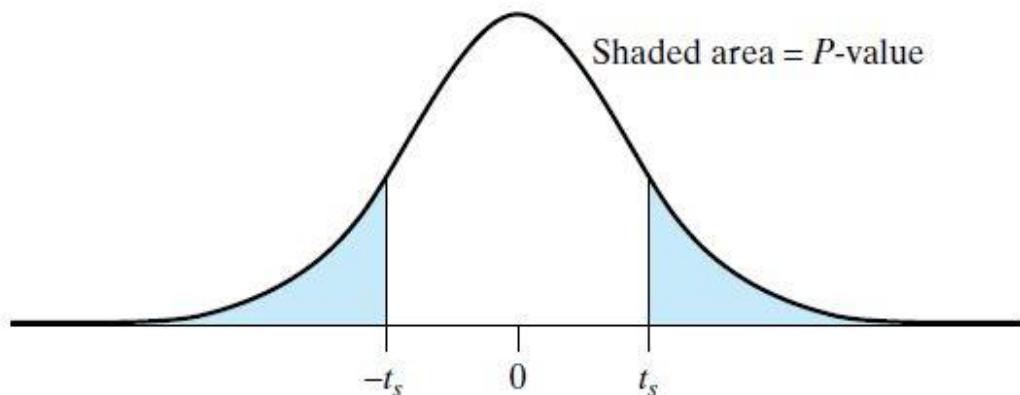
P value lies between 0 (0%) and 1 (100%) probability. The term significance is also used for this term indicated as **'sig'** in SPSS.

There are two types of significance value:

a) **True**: which is the P value what we actually get after research analysis has been done.

b) **Assumed**: is the value what we assume to get during research e.g., $P = 0.05$.

9.3: HYPOTHESIS ACCEPTANCE

9.3.1: First Rule of Thumb

When P value is close to 0 then it favors the H_1 hypothesis otherwise near 1 favors the H_o. That is why it is normally taken in between (between 0 and 1). So, the rule of thumb is: reject the H_o if P value is less than alpha (level of significance) otherwise accept H_o.

9.3.2: Second Rule of Thumb

Remember that P value can never be exact zero (0.000) and it should not be written as it is in the thesis/paper. Because there is always some value after 3 zero even up to 10 digits. For example, cruise missile (of USA) has the alpha level 17 digits after zero that is why it is better than ballistic missile and (Tomahawk Cruise Missile) is the pride of USA. But to avoid this lengthy figure, we can only write $P = 0.001$ in round off form. So, the rule of thumb is that H_1 is accepted at any level of alpha if $P = 0.000$.

9.3.3: Fixing P Value

Oldest thought was to take alpha level at 5%; however, it cannot be fixed. Remember that the range of P value depends on situation and study or researcher (always a **subjective approach**) and there is no standard for it. In case of emergency situations, it can be changed to any desirable level. For example, Sinovac (COVID-19 vaccine) was authorized for public use at the alpha level of 65%. However, Pfizer was not used in public until the clinical trial achieved alpha level of 95%. Even the P value can be 0.9 if the patient is on ventilator.

Note that the researcher P value should be less than the P value mentioned in the previous published research. For example, if you are doing a clinical trial on specific vaccine and the already published research has $P = 0.04$, then your results are only significant if their P value is less than 0.039.

9.3.4: Special Scenario

Suppose we have taken the alpha level to be 0.2 between 0 (H_1) and 1 (H_o). Now our range will vary from 0 and 0.19 and not between 0 and 1. Now the acceptance of hypothesis will depend upon P value whether it is closed to 0 or 0.19. If the P value is near 0, the H_o will be rejected otherwise H_o will be accepted. However, if the value lie in between these two (say 0.095), then we need to assume alpha two times, one above than 0.095 (10%) and one below 0.095 (9%). If the P less than alpha, then accept H_1 (or reject H_o) otherwise reject the H_1.

9.4: LEVEL OF SIGNIFICANCE

The decision as to which values go into the rejection region and which ones go into the non-rejection region is made based on the desired level of significance, designated by a. The term level of significance reflects the fact that hypothesis tests are sometimes called **significance tests**, and a computed value of the test statistic that falls in the rejection region is said to be **significant**. The level of significance, a, specifies the area under the curve of

the distribution of the test statistic that is above the values on the horizontal axis constituting the rejection region.

Rejection of null hypothesis when it is really true is known as the **Type I Error**, and acceptance of a false null hypothesis or when we fail to reject a null hypothesis that is, in fact, false is known as **Type II Error**. *The probability of type 1 error is known as the level of significance and denoted by α*, and that of type II error is generally denoted by ß. The level of significance is also known as the size of the **critical region**. If the calculated value of a test statistic lies in the critical region, the null hypothesis is said to be rejected at *α* level of significance. Generally, the level of significance depends on the objective of the study. Sometimes we may have to opt for 0.01 % or 0.001 % level of significance, particularly in relation to medical studies. A researcher has the freedom to select his or her level of significance depending upon the objective of the study.

Remember that 0.05 level of significance means that the null hypothesis may be wrongly rejected 5 times in hundred similar experiments conducted on the same population or 5% change of committing a Type I error.

9.5: TESTS IN SPSS

There are different tests used for analysis in different scenarios during research in SPSS as following:

1) Test of normality
2) Levene's test
3) Comparing Means
   a. One sample t-test
   b. Means test.
   c. Test of 2 independent means or 'Independent samples t-test' as in SPSS.
   d. Test of 2 dependent means or 'Paired samples t-test' as in SPSS.
   e. One-way ANOVA
4) General Linear Model
   a. Univariate
   b. Multivariate
   c. Repeated measures ANOVA
5) Non-parametric Tests
   a. 1-Sample K-S test
   b. 2 Independent Samples
   c. K Independent Samples
   d. 2 Related Samples
   e. K Related Samples

9.5.1: Introduction to Test of Normality

A 'test of normality' is used to check the normality of data. One of the assumptions in using all parametric tests is that the data should come from a normal population. Although the normality is a subjective approach, but it can be measured by two methods:

- **Graphically**: which is always a subjective approach like Q-Q plot, Histogram, Detrended Q-Q plot etc.
- **Numerically**: which is true measurement like Shapiro-Wilk and Kolmogorov-Smirnov tests.

Hypothesis for the normality is different than as we have discussed for DOE problems. Here the H_o indicates the normality and H_1 as non-normality e.g.:

All the knowledge belongs to God.

$$H_o = Chick\ weight\ falls\ in\ normal\ distribution$$

$$H_1 = Chick\ weight\ is\ not\ in\ normal\ distribution$$

In this scenario, the results should fall near the 1 (for Ho acceptance) and far from 0. If it is near 0, H1 will be accepted which means the data is non-normal.

9.5.2: Introduction to Levene's Test

To test the assumption of **homogeneity of variances**, SPSS computes Levene's statistic, which can be requested using the General Linear Model (GLM) command. It's a very simple and elegant test that works by doing a one-way ANOVA conducted on the deviation scores; that is, the absolute difference between each score and the mean of the group from which it came.

The hypotheses are tested as:

$$H_0 = Assume\ that\ variances\ are\ equal$$

$$H_1 = Variances\ are\ not\ equal$$

If, however, Levene's test is non-significant (i.e., $p > 0.05$) then the variances are roughly equal, and the assumption is tenable.

9.5.2.1 SPSS Procedure

We can get Levene's test using the **Explore** menu that we used in the previous section (Figure 9.50). Once the data is loaded, use following command to open the dialog box.

$$Analyze \rightarrow Descriptive\ Statistics \rightarrow Explore$$
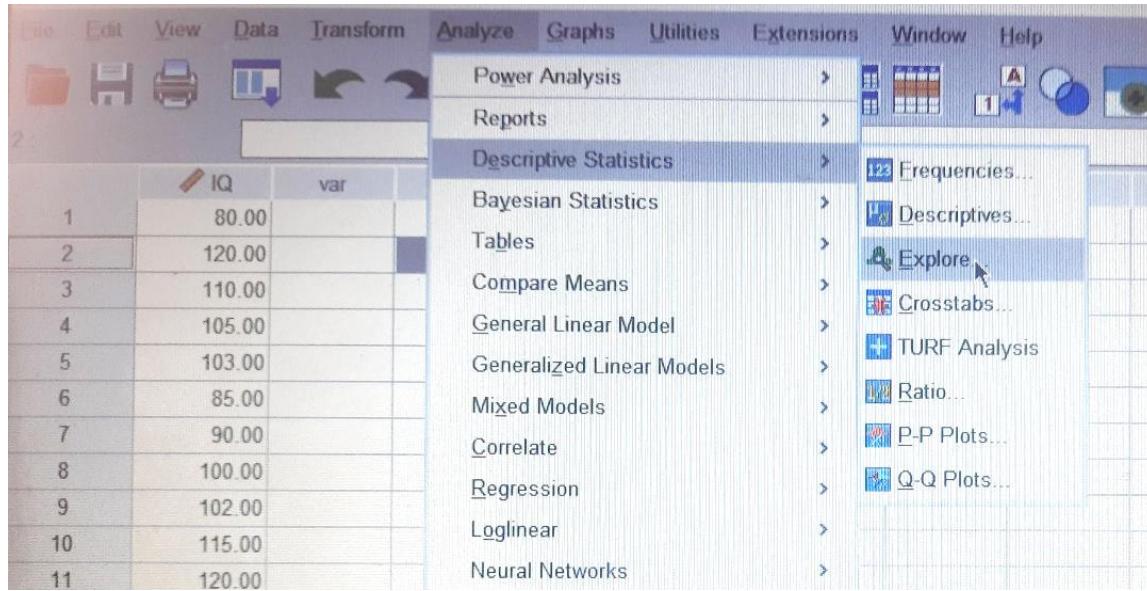
Figure 9.50: Levene's test procedure in SPSS.

To keep things simple, we'll just look at the SPSS IQ scores from this file, so transfer this variable from the list on the left-hand side to the box labelled **Dependent List** by clicking on the next to this box (Figure 9.51). Because we want to split the output by

All the knowledge belongs to God.

the grouping variable to compare the variances, select the variable Gender and transfer it to the box labelled **Factor List** by clicking on the appropriate.
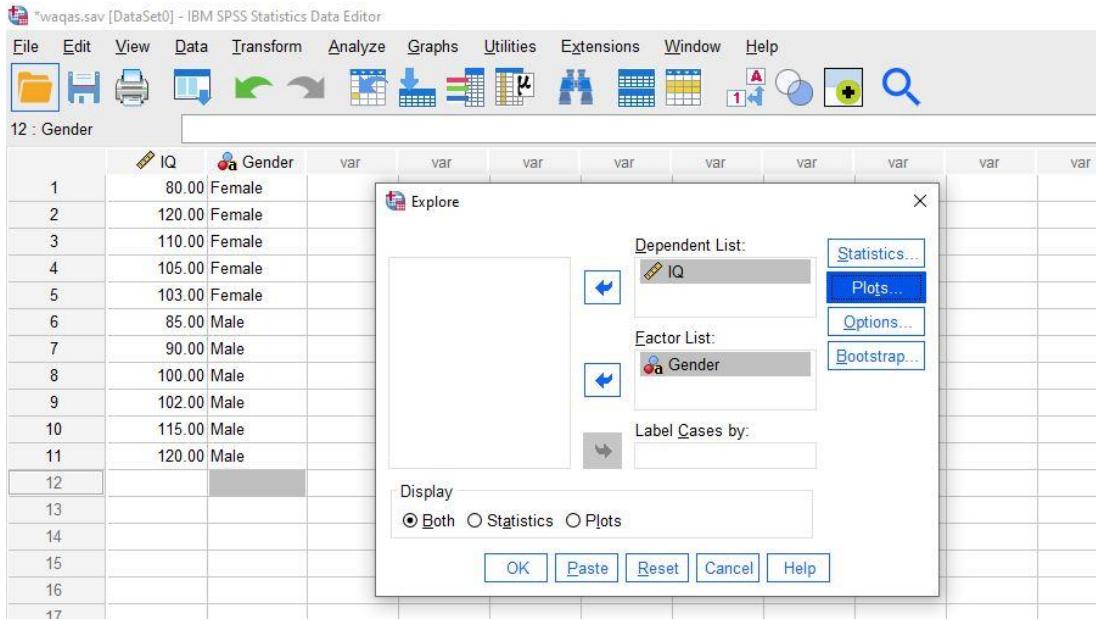
Figure 9.51: Levene's test factor selection (SPSS).

Without the presence of factor, this test can also be applied on simple raw data. Then click on **Plots** to open the other dialog box in Figure. To get Levene's test we need to select one of the options where it says **Spread vs. level with Levene's test** (Figure 9.52). If you select Levene's test is carried out on the raw data (a good place to start).
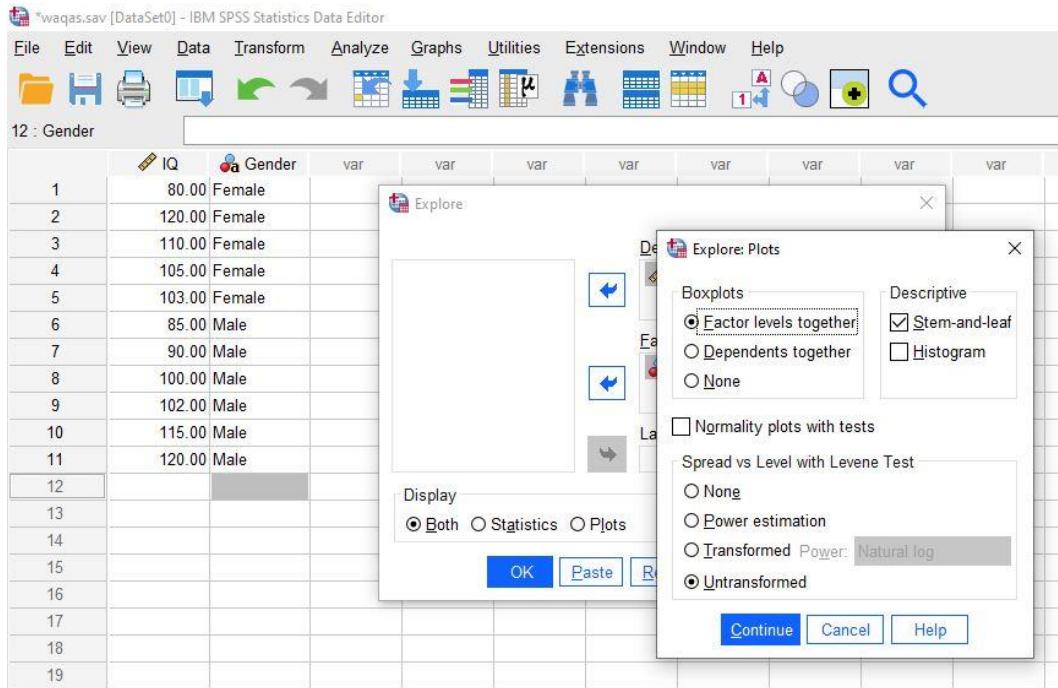
Figure 9.52: Levene's test plot selection procedure (SPSS).

All the knowledge belongs to God.

When you've finished with this dialog box click on to return to the main **Explore** dialog box and then click **OK** to run the analysis. SPSS output shows the table for Levene's test as following:

				Kurtosis			2.171	2.000
	Male	Mean					102.0000	5.56776
		95% Confidence Interval for Mean		Lower Bound			87.6876	
				Upper Bound			116.3124	
		5% Trimmed Mean					101.9444	
		Median					101.0000	
		Variance					186.000	
		Std. Deviation					13.63818	
		Minimum					85.00	
		Maximum					120.00	
		Range					35.00	
		Interquartile Range					27.50	
		Skewness					.163	.845
		Kurtosis					-1.438	1.741

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
IQ	Based on Mean	.015	1	9	.905
	Based on Median	.029	1	9	.868
	Based on Median and with adjusted df	.029	1	8.226	.868
	Based on trimmed mean	.019	1	9	.894

Figure 9.53: Levene's test output (SPSS).

You should read the statistics based on the mean. Levene's test is non-significant for the SPSS exam scores (values in the **Sig**. 0.905), indicating that the variances are not significantly different (i.e., they are similar, and the homogeneity of variance assumption is tenable).

9.5.2.2   Homogeneity Example

Following figure shows the number of hours that each person had ringing in their ears after each concert (each person is represented by a circle) (Figure 9.54). The horizontal lines represent the average number of hours that there was ringing in the ears after each concert and these means are connected by a line so that we can see the general trend of the data. Remember that for each concert, the circles are the scores from which the mean is calculated. Now, we can see in both graphs that the means increase as the people go to more concerts.

All the knowledge belongs to God.
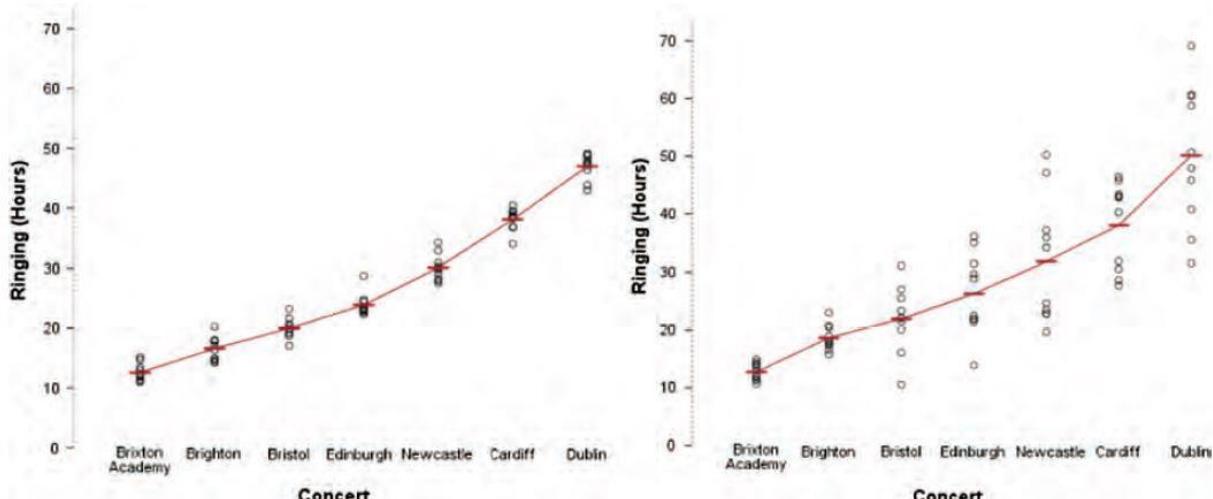
Figure 9.54: Homogeneity and heterogeneity of variances.

So, after the first concert their ears ring for about 12 hours, but after the second they ring for about 15–20 hours, and by the final night of the tour, they ring for about 45–50 hours (2 days). So, there is a cumulative effect of the concerts on ringing in the ears. This pattern is found in both graphs; the difference between the graphs is not in terms of the means (which are roughly the same), but in terms of the spread of scores around the mean. If you look at the left-hand graph the spread of scores around the mean stays the same after each concert (the scores are fairly tightly packed around the mean).

Put another way, if you measured the vertical distance between the lowest score and the highest score after the Brixton concert, and then did the same after the other concerts, all of these distances would be fairly similar. Although the means increase, the spread of scores for hearing loss is the same at each level of the concert variable (the spread of scores is the same after Brixton, Brighton, Bristol, Edinburgh, Newcastle, Cardiff and Dublin). This is what we mean by homogeneity of variance. The right-hand graph shows a different picture: if you look at the spread of scores after the Brixton concert, they are quite tightly packed around the mean (the vertical distance from the lowest score to the highest score is small), but after the Dublin show (for example) the scores are very spread out around the mean (the vertical distance from the lowest score to the highest score is large). This is an example of heterogeneity of variance: that is, at some levels of the concert variable the variance of scores is different to other levels (graphically, the vertical distance from the lowest to highest score is different after different concerts).

9.5.3: Compare Means

9.5.3.1 Introduction to One Sample t-Test

We use this test to check the mean of one sample with a hypothesized number. For example, if we have average weight of 20 chickens to be hypothesized as 1500g (null hypothesis) and we want to actually check whether in real world it is the same or not (alternative hypothesis). For this purpose, **One Sample t-test** is used.

9.5.3.2 Introduction to Means Test

It is used for the test of means (simply give means along with other statistical parameters when selected e.g., standard deviations, median, standard error etc.). It is actually a z-test. No hypothesis testing is possible in this test.

All the knowledge belongs to God.

9.5.3.3 Introduction to Independent-Samples T Test

In this test, we have two conditions, but different subjects participate in each condition. One group participate in one condition and other in second condition e.g., means of male and female IQs. That is why **'Grouping of Variables'** is required to run this test. Thus, it can be used to test the hypothesis whether means are same or different/which is better and for this purpose **Independent Samples t-test** is used.

9.5.3.4 Introduction to Paired-Samples T Test

In a paired design, the observations occur in pairs; the observational units in a pair are linked in some way, so they have more in common with each other than with members of another pair. This test is used when the means of data are somehow related to a single experimenter or two different experimenters are working on the same data. This test is also known as matched-pairs or paired- matches t-test. For this purpose, Paired Samples t-test is used. It is different from the previous test as there is a correlation between two classes here. It can be easily understood by different scenarios as following:

- Suppose we have class of 20 people, 10 male and 10 female but all are married. And we ask them the age of marriage. Their answer is in two groups: male and female. Now we have two means, but they are different. So, their means are independent as the genders are different (independent samples t-test). Now suppose, we have 10 couples (20 people) and we ask them the same question. Although their means of ages will be different but they are dependent due to people being couples (related to each other). Here we will apply test for paired samples t-test.
- If a drug is tested for lowering blood pressure for different times on the same patient (say for 3 months in 2022 and 3 months in 2023), the means will be paired and the test will be paired samples t-test.
- If two levels of feeds are given to same chicken (say one feed for first 2 weeks and then other feed for nest 2 weeks) their means will be paired due to association with the same chicken. Here means are paired and paired samples t-test will be applied.

Remember that if one person does two different things or one task is done by two different persons, it will be paired test (paired means).

9.5.3.5 Introduction to One-way ANOVA

It is the extension of Student's T-test to more than 2 groups. In fact if you do one-way ANOVA on 2 groups, it will give the same results as Student's T-test for independent samples. It is execution of CRD in fact. If the effect of only one factor on some dependent variable is investigated then it is One-way ANOVA. For two factors (RCBD or 2 factor factorial) simultaneously, it will be Two-way ANOVA.

9.5.4: General Linear Model (ANOVA Measures)

Analysis of variance is a technique whereby the total variation present in a set of data is partitioned into two or more components. Associated with each of these components is a specific source of variation (factor), so that in the analysis it is possible to ascertain the magnitude of the contributions of each of these sources to the total variation. It can be used to measure the means of 3 or more groups. In this analysis, the null hypothesis of no difference among the group means is tested against the alternative hypothesis that at least one group mean differs. Depending upon the way of treatment (independent variables) are allocated, to the subject, ANOVA is classified into 3 categories:

- Independent measures ANOVA
- Repeated Measures ANOVA

All the knowledge belongs to God.

- Mixed ANOVA.

Only the first two will be discussed here.

9.5.4.1   Test for >2 Independent Means

If each subject in a study receives only one treatment, then such studies are investigated by using One-way Independent measures ANOVA. It is also known as **between-group design** or **between subjects.** There are normally two types of tests for this measure in SPSS:

- Univariate: it may be used for
    - One-way ANOVA
    - Two-way ANOVA
    - ANCOVA (analysis of covariance)
- Multivariate: it can be used for multi-way ANOVA or MANOVA

9.5.4.2   Test for >2 Dependent Means or Repeated Measures ANOVA

Repeated measures is when same subject participate in all the conditions of an experiment e.g., we can have the effect of feed (with 5 protein levels) on weight gain of chicken in different time period of its age (on weekly basis) or students want to check their marks consistency by submitting same papers to 4 different teachers.

If each subject in a study receives all the treatments, then it is **repeated measures ANOVA**. It is also known as **within-subjects design**. That treatment is known as **within-subjects factor**. The same subject participates in both conditions. For example, effect of feed on weight gain of a group of chickens 5 weeks before and after the feed is given to them. Now the subjects (chicken) are the same. Similarly, if we take body weight of chicks taken every week up to 5 weeks. Now we will have 5 means of weight (more than 2 means) repeated each week. This scenario is suitable for repeated measure ANOVA (RMA).

This design has several advantages. It reduces the unsystematic variation and is economical (fewer subjects are required). However, the independence of data (means coming from the different subjects) is violated. Remember that even if we have many means and only one factor repeats among those (dependent), and rest of factors remain independent, we will apply RMA.

9.5.5: Introduction to Nonparametric Tests

9.5.5.1   1 Sample K-S Test

Following are the tests used in this section for analysis:

- 1 Sample K-S Test: is an alternative version of parametric test for One-Sample T Test.
- Independent Samples Test: is an alternative version of parametric test for Independent Samples T Test.
- 2 K Independent Samples Test: is an alternative version of parametric test for One-Way ANOVA.
- 2 Related Samples Test: is an alternative version of parametric test for Paired Samples T Test.
- 2 K Related Samples Test: is an alternative version of parametric test for One-Way Repeated Measures ANOVA.

All the knowledge belongs to God.

CHAPTER 10: TEST OF NORMALITY

10.1: INTRODUCTION

We come now to the most important distribution in all of statistics—the normal distribution. The formula for this distribution was first published by **Abraham De Moivre** (1667–1754) on November 12, 1733. This test helps the researcher to check the normality of data and how to present/make data normal (even if it is abnormal) using SPSS**.** We can convert abnormal data into normal without resorting to illegal means.

Normality can be checked by testing the significance of kurtosis and skewness. If these two statistics are insignificant, the data is normal. In case of these two being insignificant, **Shapiro-Wilk** and **Kolmogorov-Smirnov** tests are used to test the normality using SPSS. ***Kurtosis is a measure of the degree to which a distribution is "peaked" or flat in comparison to a normal distribution whose graph is characterized by a bell-shaped appearance.*** Kurtosis is also known as **pointiness**.

10.1.1: Shapiro–Wilk and Kolmogorov–Smirnov Test

The Shapiro–Wilk test is more suitable for testing normality in case of small sample (N = 50), but it can be used for the sample size up to 2000. However, if the sample size is large, then the Kolmogorov–Smirnov test is used for checking normality of data. Later test gets its name from A. Kolmogorov and N. V. Smirnov, two Russian mathematicians who introduced two closely related tests in the 1930s.

One of the limitations of these tests is that in case of large sample, you are more likely to get significant results. In other words, in large sample these tests become significant even for slight deviation from normality. While testing the normality with SPSS, one can also choose the option for identifying outliers using boxplot.

10.2: NORMALITY IN SPSS

Suppose we have IQs range of students (Table 10.1) and we want to check normality of the data.

Table 10.1: IQ of students.

113	84	81	90	99
85	100	120	82	113
88	110	95	105	93
87	116	107	91	96

10.2.1: SPSS Procedure

The hypothesis will be:

$$H_1 = The\ IQ\ data\ of\ students\ is\ not\ normally\ distributed$$

The IQ data of student is first entered into **Data View** tab of SPSS after defining the data set into **Variable View** as following:
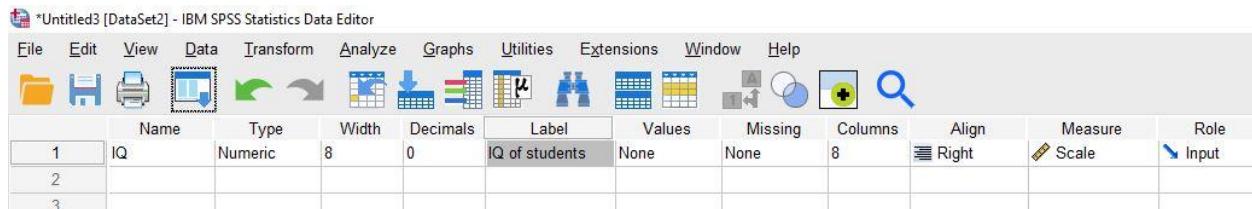
Figure 10.1: Defining the IQ variable in Variable View (Test of Normality).

The **Data View** will look like as following:

All the knowledge belongs to God.

Figure 10.2: Data input for Test of Normality (SPSS).

After preparing the data file, follow the below mentioned sequence of commands as shown in the Figure 10.3:

Analyze → Descriptive statistics → Explore
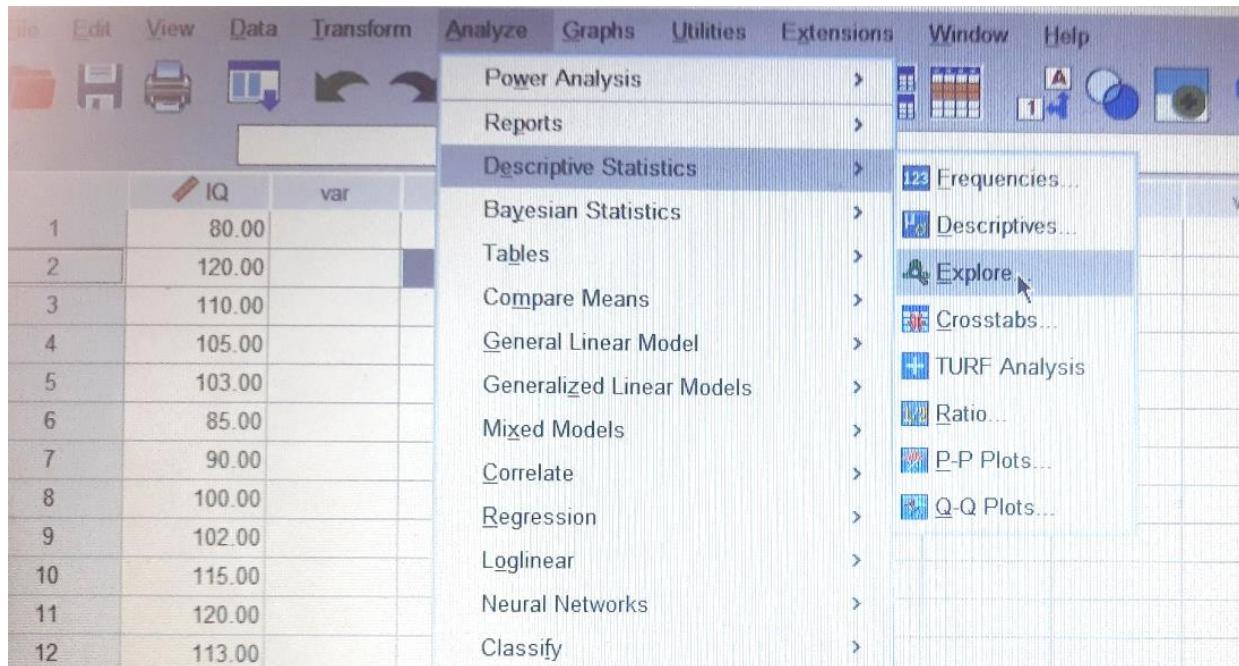
Figure 10.3: Test of normality command sequence (SPSS).

After clicking **Explore**, select the variables for testing normality and identifying outliers (Figure 10.4). Select the variable and shift it/them into **Dependent List** section.
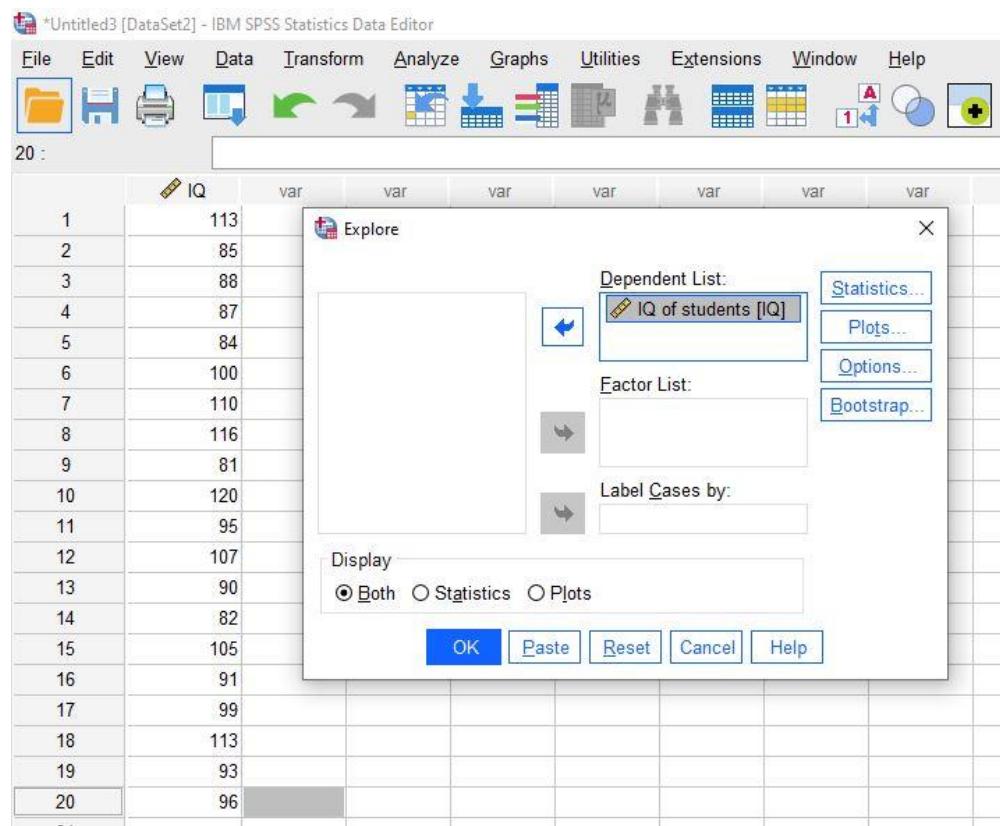
Figure 10.4: Explore command inputs (Test of Normality).

All the knowledge belongs to God.

Click on the **Statistics** commands and select the **Outliers** option (Figure 10.5). You can change the **Confidence Interval for Mean** whatever you want.
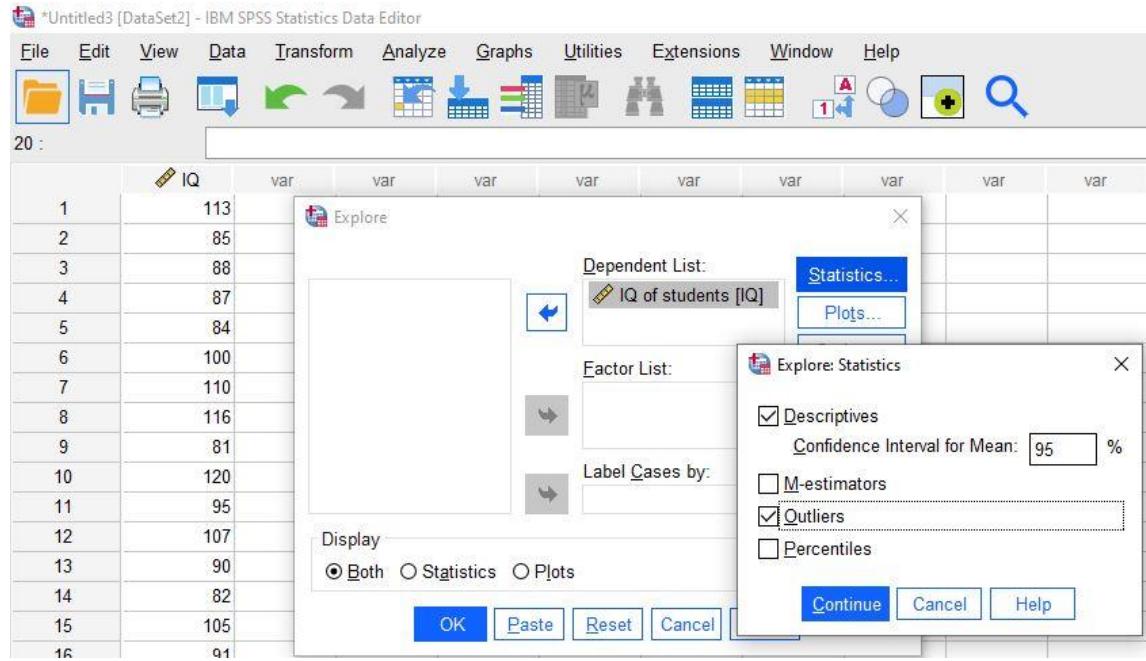
Figure 10.5:Statistics option editing (Test of Normality).

Click on **Continue**, then on **Plot** command in the same screen and select the **Normality plots with test** option (Figure 10.6). It'll generate the output of Shapiro-Wilk test and Q-Q plot. **Histogram** can also be checked if required.
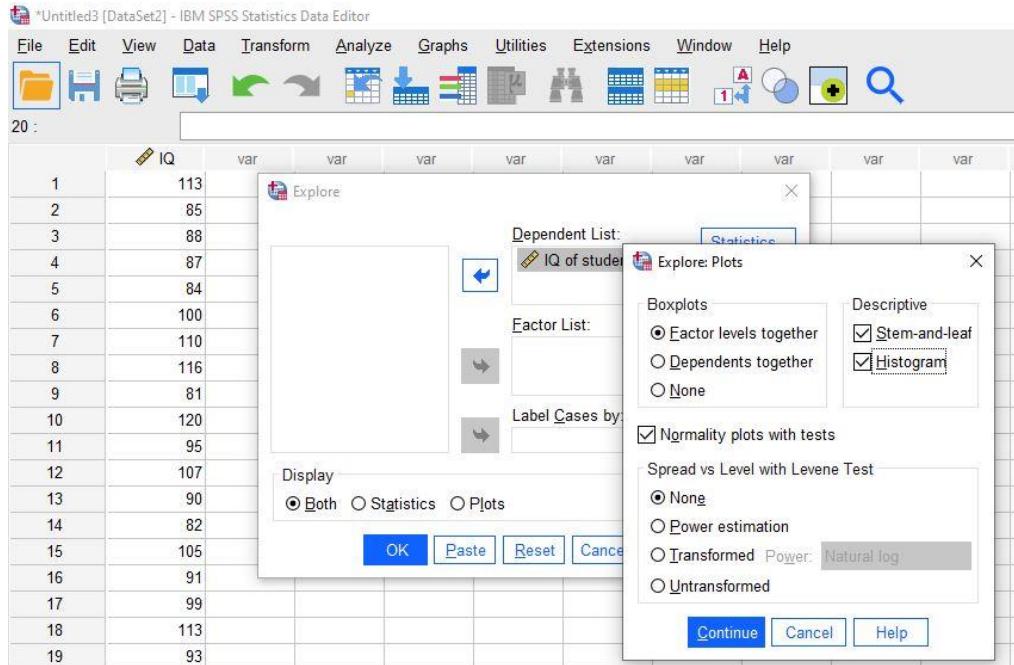
Figure 10.6: Plot command editing (Test of Normality).

All the knowledge belongs to God.

Let the other options remain selected by default. Click on the **Continue** and **OK** options (Figure 10.7) to get outputs.

Figure 10.7: Test of Normality output (Descriptives).

10.2.2: SPSS Output

First table in Figure 10.7 is **Case Processing Summary** which shows that all the data has been shown (20 inputs) and nothing is missing.

In **Descriptives** table, there is plenty of information available for researcher:

- **Mean** with **Std Error** (standard error) is shown. Remember that less **Std Error** shows that the data is normal.
- The **95% Confidence Interval for Mean**: shows the upper and lower limit (as **Upper Bound** and **Lower Bound,** respectively) of the 95% limit set for data which is 92-103 in this case.
- **5% Trimmed Mean**: is actually a rough mean estimate of data if the data has been trimmed (cut) 5% above and below the mean. Note that it is (97.4)

All the knowledge belongs to God.

very close to actual mean (97.7). This indicates that there are no/very few outliers in our data. For a set of sample data containing 'n' measurements, we calculate the 100a percent trimmed mean as follows:

- o Order the measurements.
- o Discard the smallest 100α percent and the largest 100α percent of the measurements. The recommended value of α is something between 0.1 and 0.2.
- o Compute the arithmetic mean of the remaining measurements.

- **Median**: The middle score or median is the appropriate measure of central tendency for ordinal level raw data.
- **Median**: The median of a finite set of values is that value which divides the set into two equal parts such that the number of values equal to or greater than the median is equal to the number of values equal to or less than the median. If the number of values is odd, the median will be the middle value when all values have been arranged in order of magnitude. When the number of values is even, there is no single middle value. Instead there are two middle values. In this case the median is taken to be the mean of these two middle values, when all values have been arranged in the order of their magnitudes.
- **Mode**: The mode of a set of values is that value which occurs most frequently. If all the values are different there is no mode; on the other hand, a set of values may have more than one mode. A distribution with two modes (there are two bars that are the highest), which is said to be bimodal. It's also possible to find data sets with more than two modes (multimodal).
- **Variance**: When the values of a set of observations lie close to their mean, the dispersion is less than when they are scattered over a wide range. Since this is true, it would be intuitively appealing if we could measure dispersion relative to the scatter of the values about their mean. Such a measure is realized in what is known as the variance. In computing the variance of a sample of values, for example, we subtract the mean from each of the values, square the resulting differences, and then add up the squared differences. This sum of the squared deviations of the values from their mean is divided by the sample size, minus 1, to obtain the sample variance. However, it has no significance for researcher.
- **Degrees of Freedom**: The reason for dividing by n-1 rather than n, as we might have expected, is the theoretical consideration referred to as degrees of freedom. In computing the variance, we say that we have n-1 degrees of freedom. In simple words, these are the data sets that are free to vary except one we have excluded.
- **Standard Deviation**: the average dispersion among the units in data. Our data has inputs which are at 13.2 units away from each other.
- The **Interquartile Range** (IQR): it is the difference between the third and first quartiles or cut off the top and bottom 25% of scores and calculate the range of the middle 50% of scores – known as the interquartile range. A large IQR indicates a large amount of variability among the middle 50 percent of the relevant observations, and a small IQR indicates a small amount of variability among the relevant observations. Since such statements are rather vague, it is more informative to compare the interquartile range with the range for the entire data set.

All the knowledge belongs to God.

In **Test of Normality Table** (Figure 10.8), output of two different procedures (Kolmogorov-Smirnov and Shapiro-Wilk tests) is shown, if they are significant (reject H_o), the data is non-normal. Thus, for the data to be normal these tests should be non-significant. Note that they might have very different outputs (0.200 versus 0.252).
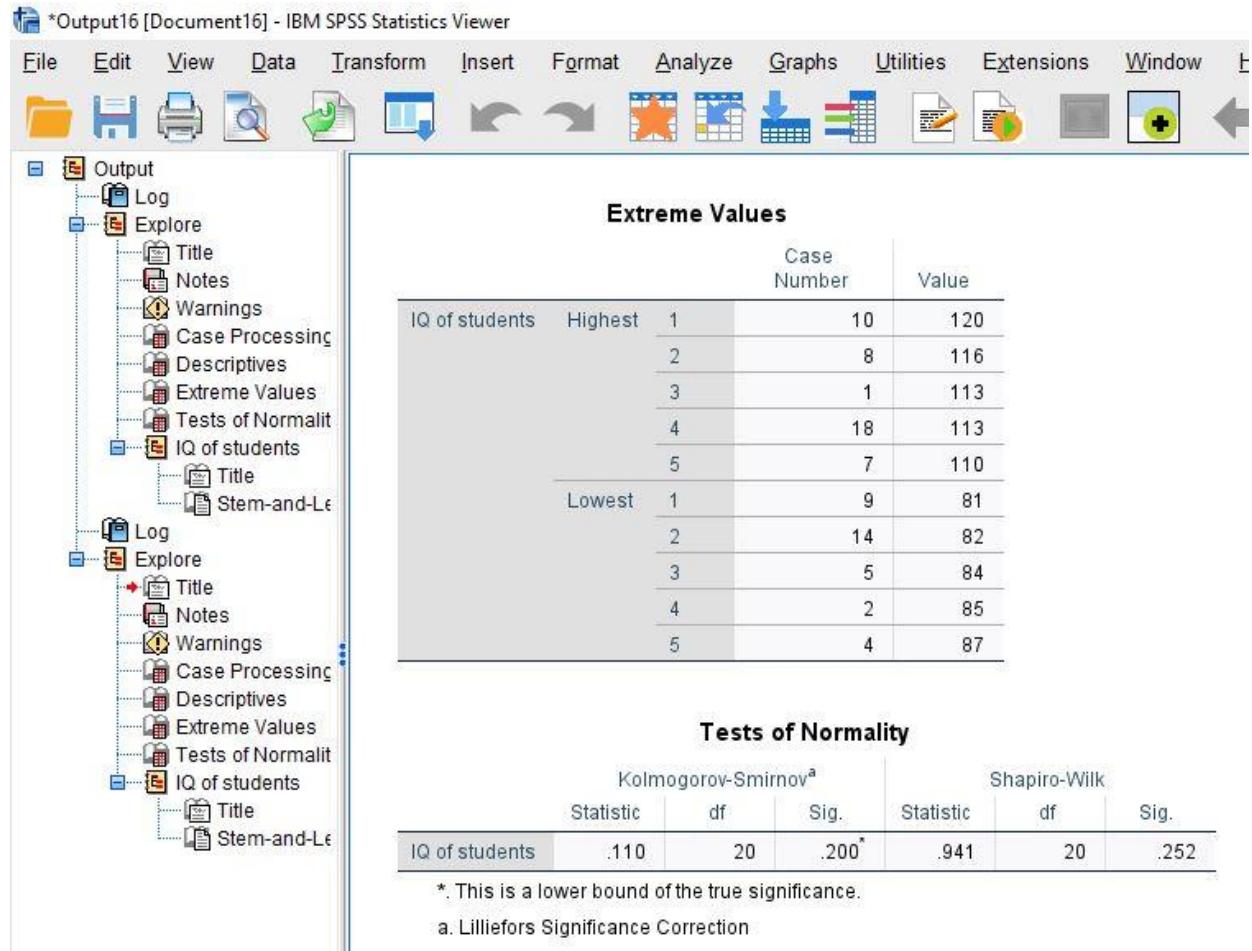
Figure 10.8: Test of Normality table (output).

When a variable can take on many values, the values in a frequency or percentage distribution become too numerous to be informative. In these cases, it can be helpful to group values together before tallying the cases. **Histograms** plot the relative frequency of cases within such groups (Figure 10.9). The Histogram command creates a grouped frequency distribution. The range of scores is split into evenly spaced groups. The midpoint of each group is plotted on the X-axis, and the Y –axis represents the number of scores for each group.

In histograms, we group categories into equal intervals. It may also indicate the normality of data if the spreadness of bars are near center. In this case it is a good option to use this histogram in thesis.

All the knowledge belongs to God.
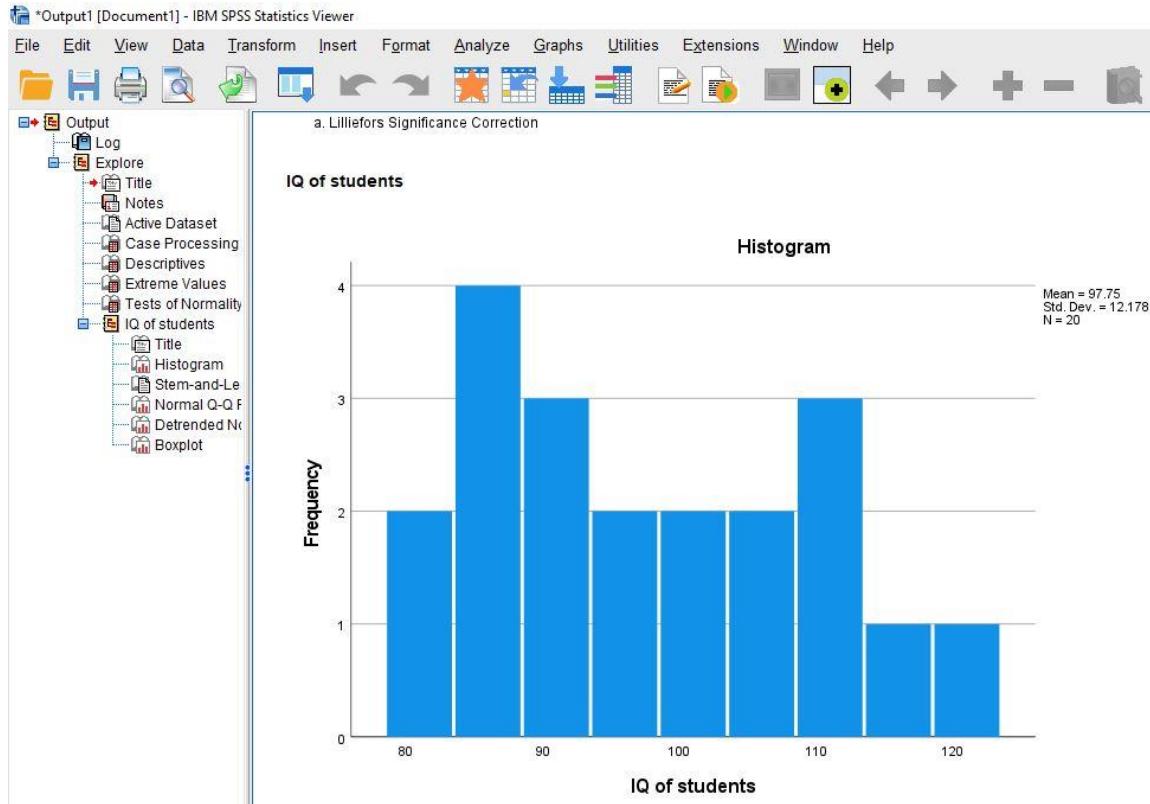
Figure 10.9: Histogram output (Test of Normality).

A **Q–Q plot** (Figure 10.10) is very similar to the P-P plot except that it plots of quantiles the data set instead of every individual score in the data. Quantiles are just values that split a data set into equal portions. In short, then, the Q–Q plot can be interpreted in the same way as a P–P plot but it will have less points on it because rather than plotting every single data point it plots only values that divide the data into equal parts (so, they can be easier to interpret if you have a lot of scores).

Q–Q plot is a graphical way of checking the normality of data. It compares the two probability distributions by plotting their Quantiles against each other. If distribution of sample data are similar to that of standard normal distribution, all the points in the Q–Q plot will lie very close to the line. Or if the data are normally distributed, then the observed values (the dots on the chart) should fall exactly along the straight line (meaning that the observed values are the same as you would expect to get from a normally distributed data set).

All the knowledge belongs to God.
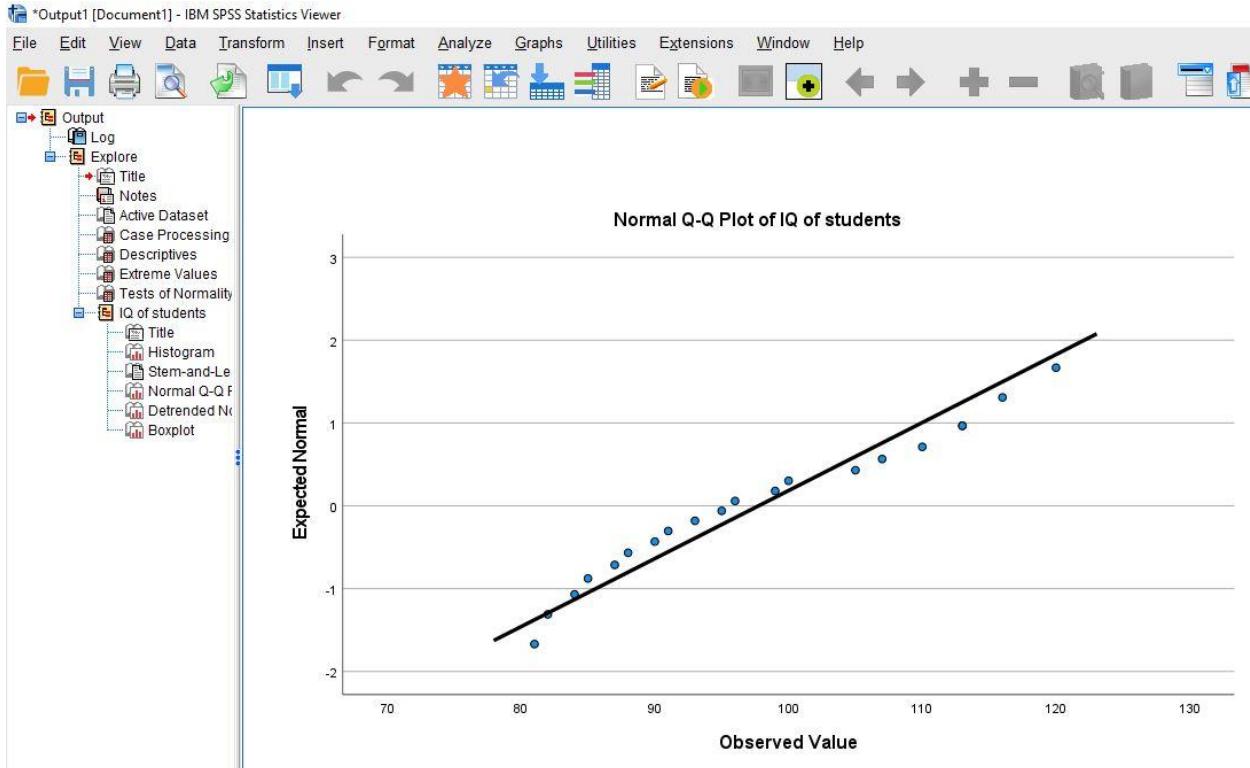
Figure 10.10: Q-Q plot output (Test of Normality).

Detrended Normal Q-Q Plot indicates the data above and below the central line (Figure 10.11). If the line is in between the upper and lower range (in other words, if dots are symmetrical in quantity at both sides), the data is normal.
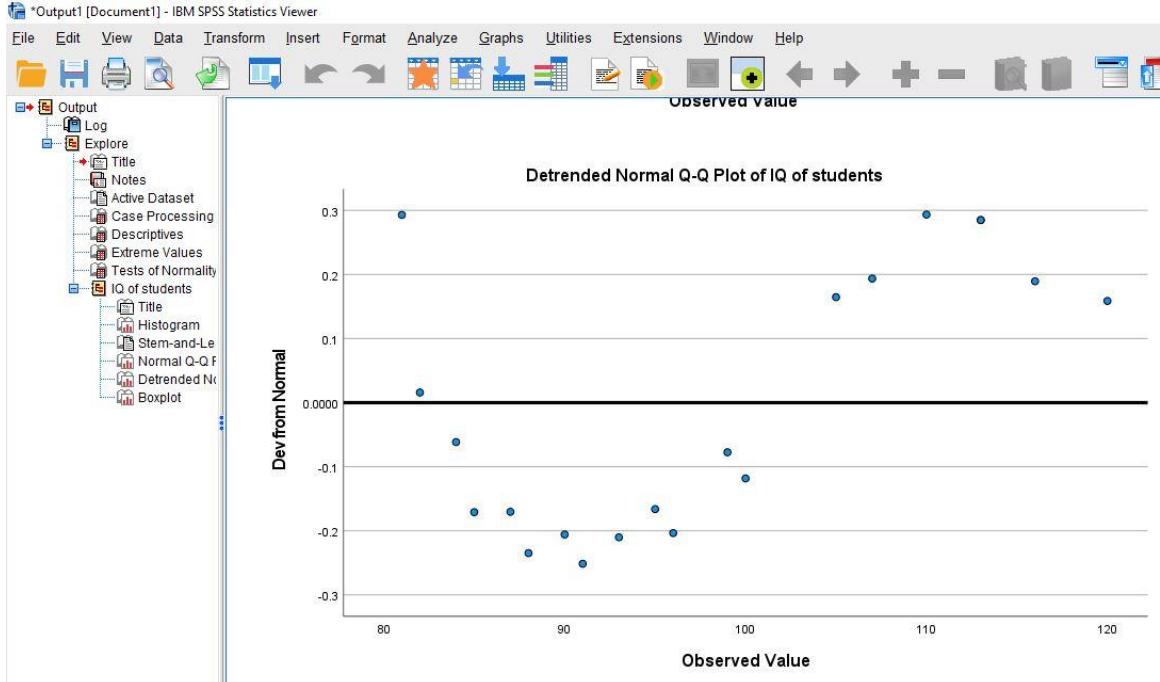
Figure 10.11: Detrended Q-Q plot output (Test of Normality).

All the knowledge belongs to God.

A useful visual device for communicating the information contained in a data set is the **box-and-whisker plot** or **Boxplot**. The construction of a box-and-whisker plot (sometimes called, simply, a boxplot) makes use of the quartiles of a data set. The box plot was popularized by **John Tukey** and helps us visualize several key statistical features of a variable. Figure 10.12 shows the basic features of the original boxplot as laid out by Tukey (in 1977).
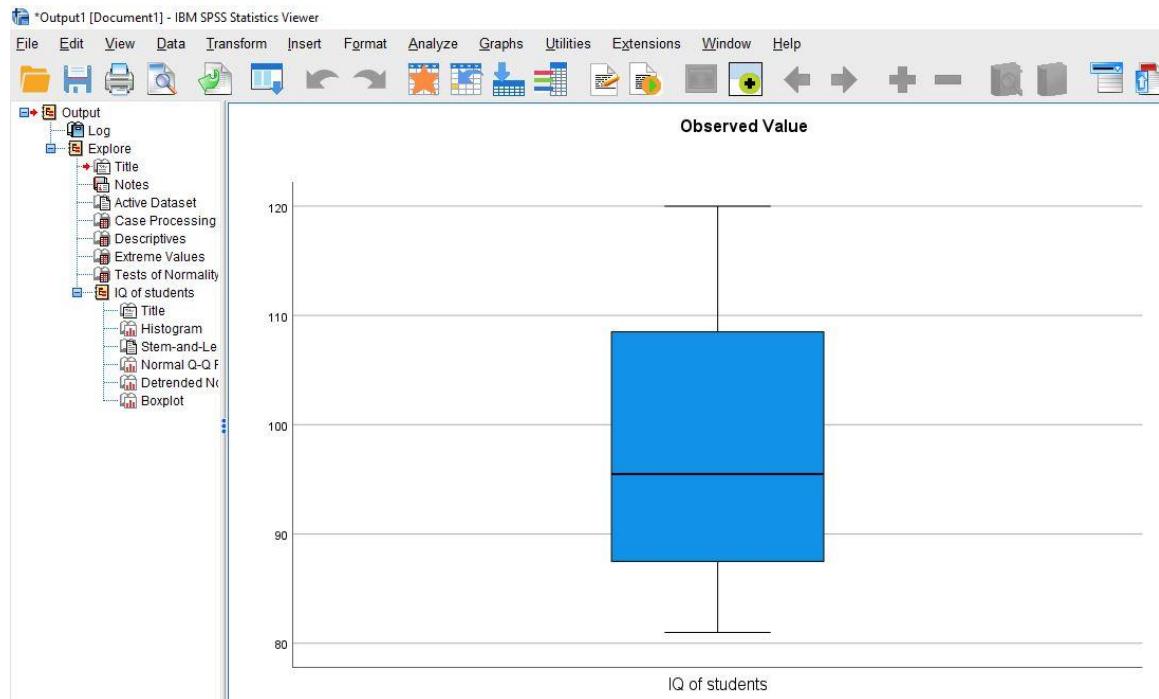
Figure 10.12: Boxplot output (Test of Normality).

The line between shaded area (blue) should be in center of area for the data to be normal. This plot can also be used in thesis. The outlier is an unusual data which a researcher tries to remove from the sample if the results drawn from the sample need to be generalized for the population of interest. If the researcher feels that the data is genuine, then he may decide to keep it in the study. Figure 10.12 shows the Boxplot for variable. It describes the distribution of data and identifies outliers if any. Usually, any data outside the mean ± 2SD is taken as outlier. The SPSS computes outlier based on interquartile range. However, you may keep the data in your sample even if it lies outside this range provided you are convinced that such score is genuine and can be obtained by the subjects easily. It can be seen in Figure 10.12 that the no score is shown as outlier.

As its name implies, a box is a key feature. When displayed vertically, the box is drawn so that its upper and lower borders are the 3rd and 1st quartiles respectively (the 75th and 25th percentiles) and a horizontal line is drawn within the interior of the box at the median (50th percentile). In the standard box plot, lines are drawn from the box boundaries to the extreme values (the minimum and maximum) of the variable (Figure 10.13). Tukey conceived of this as a 5-number summary of the data (the two extremes, the median, and the 1st and 3rd quartiles).

All the knowledge belongs to God.
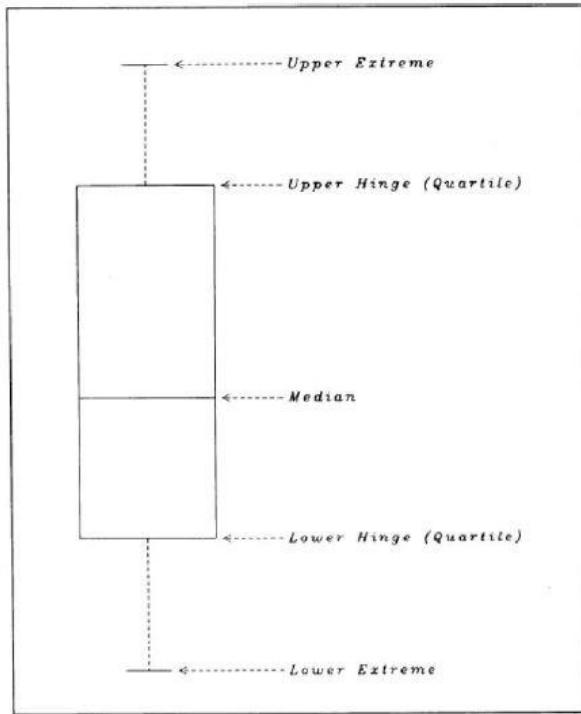
Figure 10.13: Explanation of boxplot (Test of Normality).

At the center of the plot is the median, which is surrounded by a box the top and bottom of which are the limits within which the middle 50% of observations fall (the interquartile range). Sticking out of the top and bottom of the box are two whiskers which extend to the most and least extreme scores respectively. Or the lines extending from the box are referred to as whiskers and are used to help identify values that fall far from the rest of the data. Such values are often referred to as outliers. In the context of a box plot, an outlying value is defined based on the interquartile range. *If a value falls more than 1.5 times the interquartile range above the 3rd quartile or more than 1.5 times the interquartile range below the 1st quartile then it is an outlying value*. Typically, the whiskers are drawn to the closest value falling within these bounds. Then, outlying values are marked where they fall past these bounds.

Remember that only show the plot which indicates normality for all the variables in your thesis. You cannot use different kinds of plots in a thesis.

CHAPTER 11: COMPARING MEANS IN SPSS

11.1: INTRODUCTION

In comparative studies, we intend to compare the means of two groups and our focus is to test whether the difference between the two groups is significant.

11.2: MEANS TEST

This test includes the z-test for comparing two population means (μ) to each other as well as comparing a sample mean (X) to the parent population mean (μ), assuming that the parent population is normally distributed.

In comparative study, we intend to compare means of multiple groups e.g. to compare IQ of boys and girls. For this purpose, two statistical tests, t-test (for small samples: <30) and z-test (for large samples: >30) are used. Remember that we cannot do hypothesis testing using Means test.

11.2.1: SPSS Procedure

Suppose we have running data of Nasir and Khan (Table 11.1) and we want to compare their means:

Table 11.1: Running data of Khan and Nasir.

Khan	120	110	99	100	140	133	130	140	144	138
Nasir	100	100	98	95	88	88	88	85	90	95

First define the two variables in **Variable View** tab as following:
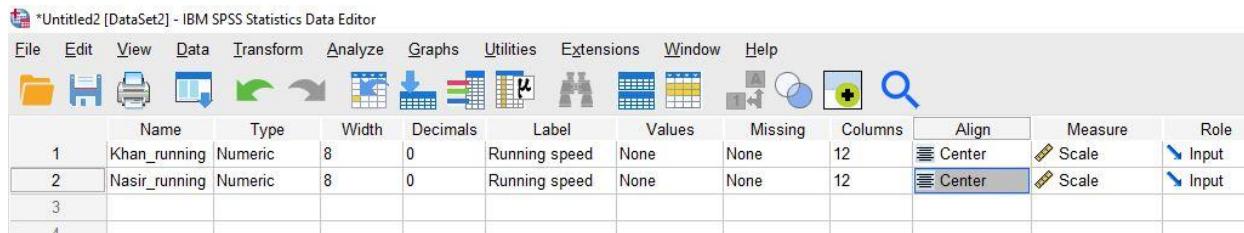
Figure 11.1: Defining variables in Variable View (Means Test).

Insert data into SPSS in two columns and apply **Means** test using commands:
$$Analyze \rightarrow Compare\ Means \rightarrow Means$$

All the knowledge belongs to God.
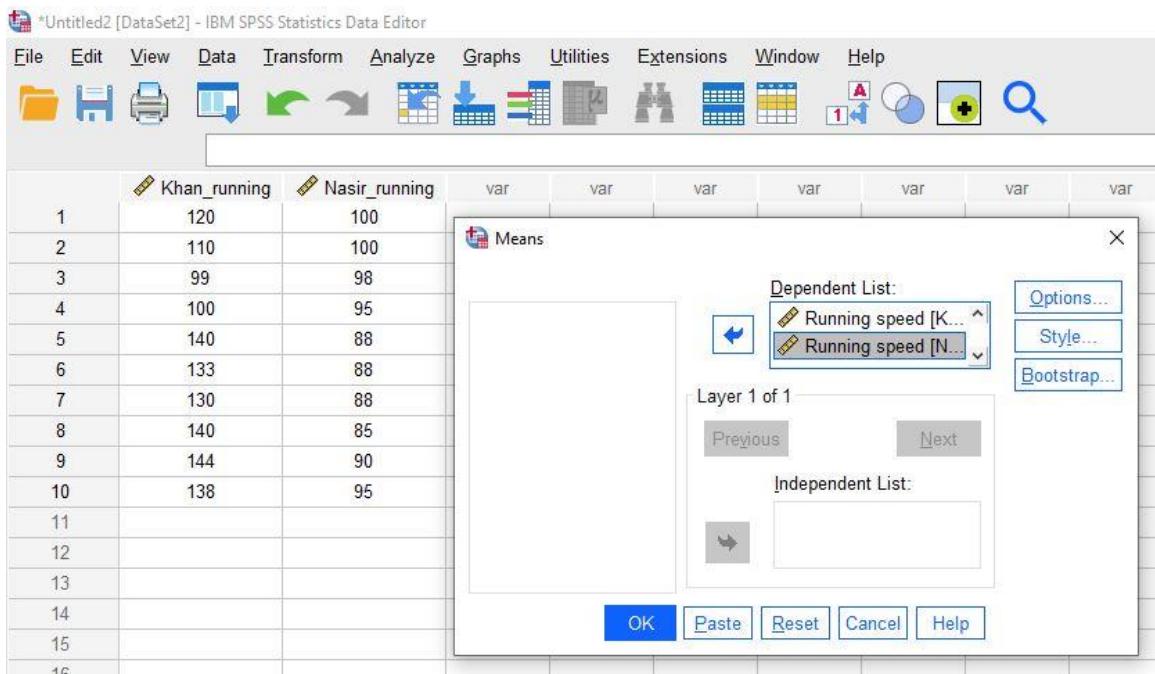
Figure 11.2; Data inputs (Means test).

Remember that we can use **Layer** option to sort out means into individual subsets i.e., female and male. We can insert different parameters to be checked using **Option** tab i.e., std error or median etc., as following:
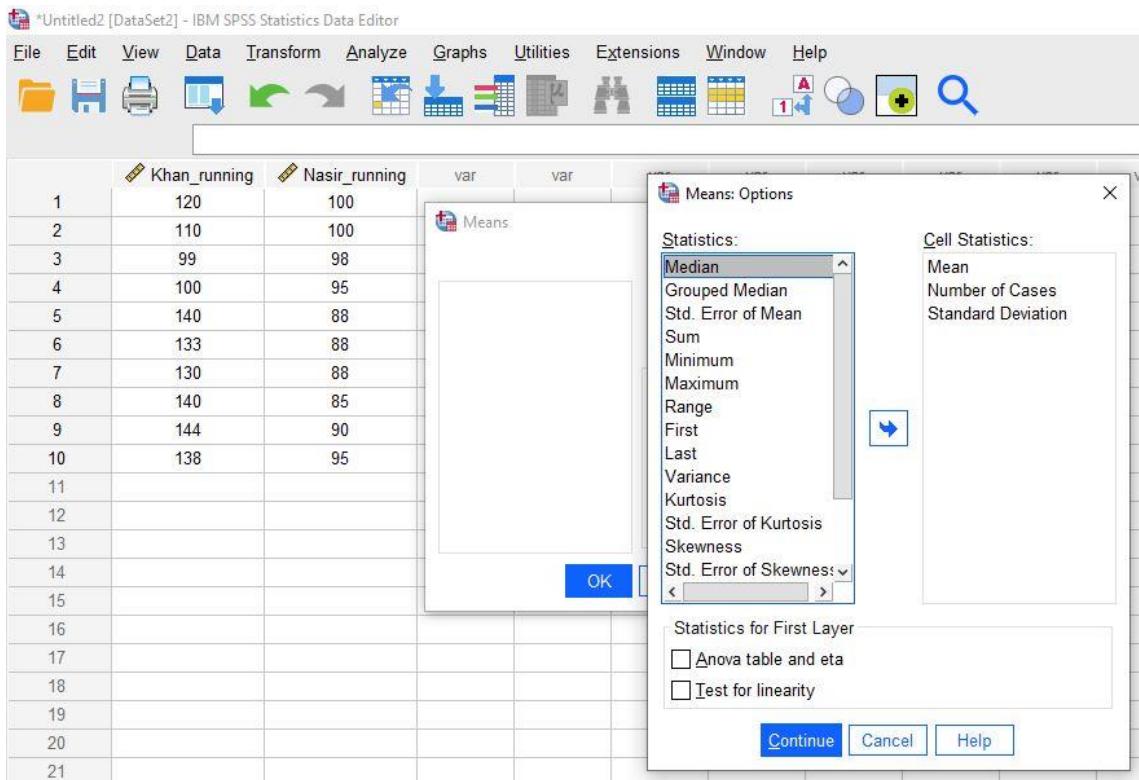
Figure 11.3: Options selection (Means test).

All the knowledge belongs to God.

After clicking **Continue**, we have output:
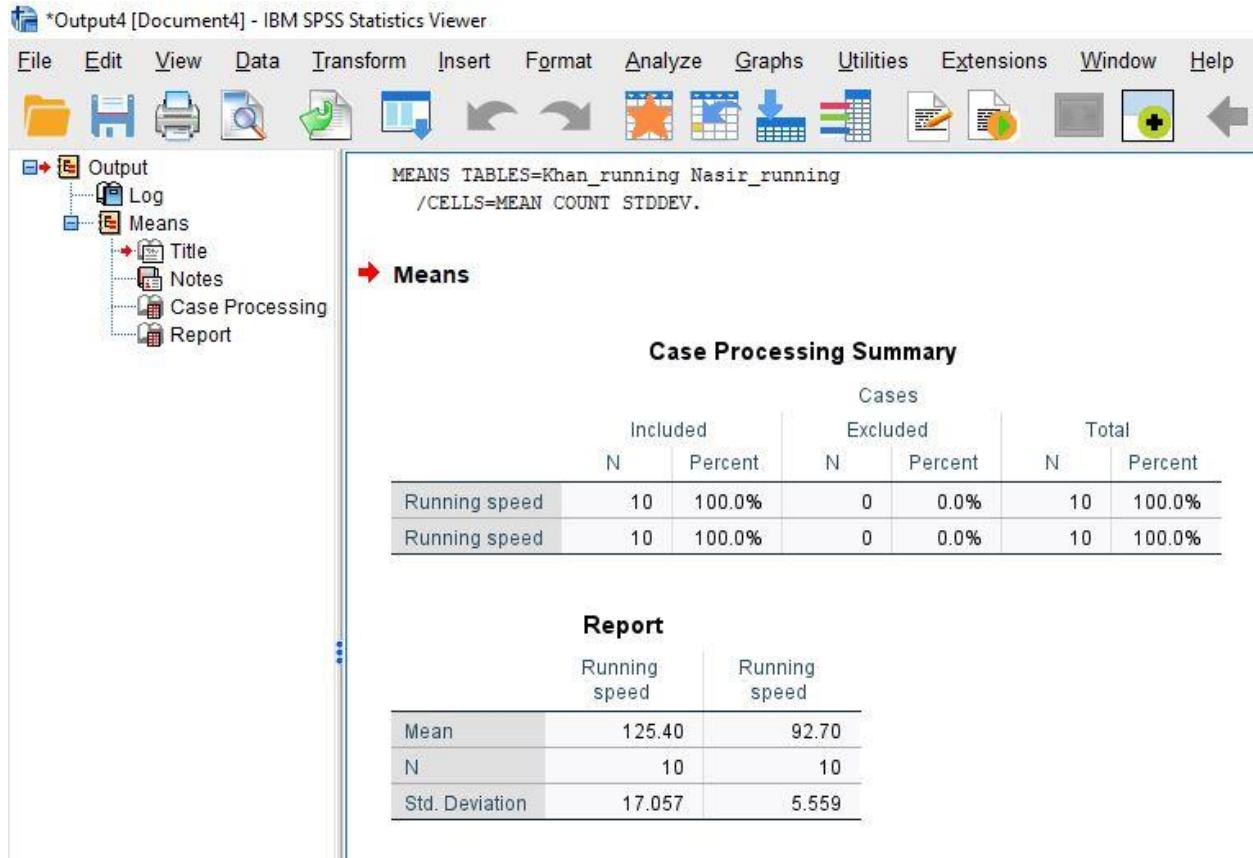
Figure 11.4: Means test output.

Note that we don't have any significance value. We just have means and other parameters which we have selected (std. error and std. deviation).

11.3: ONE SAMPLE T-TEST

It is used for comparing the mean value of a sample to a predefined population mean. It is assumed that the population mean is known in advanced. In other words, this test is used when we have only one group of data but hypothesized. It is useful for determining if the current set of data has changed from a long-term value (e.g., comparing the current year's temperatures to a historical average to determine if global warming is occurring).

For example, a specific protocol of exercise used in a physical therapy center brings relief to the spondylitis patients within a 20-day session. When introducing a new set of exercise, it is administered to 25 patients, and the days until the exercise shows an effect are recorded. The mean days of getting the relief is 16 days with a standard deviation of 4 days. Can it be concluded that the new exercise reduces the time until a patient receives relief from spondylitis pain? In this example, sample mean (the mean relief days for spondylitis patients) requires to be compared with a predefined limit. Here, the limit is fixed (16 days) and well known in advance.

11.3.1: SPSS Procedure

Suppose we have data of Khan's average running score per day (Table 11.2) and we hypothesized that his average is 20km/hour. Check whether our hypothesis is true or not?

All the knowledge belongs to God.

Table 11.2: Khan's running speed.

Khan's Running	15	20	25	25	25	23	22	19	25	22

We shall have hypothesis:

$H_1 = The\ average\ running\ of\ Khan\ is\ different\ from\ 20\ km\ per\ hour$

The variables included are:
- $Test\ value = 20$
- $Dependent\ variable = Khan's\ running.$

First, define the variable in **Variable View** as following:
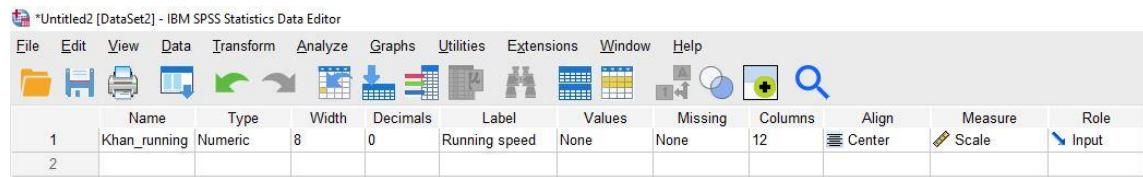
Figure 11.5: Defining variable in Variable View (One sample T Test).

After entering all the data in **Data View**, the following steps should be followed for computing t-statistic:

$Analyze \rightarrow Compare\ Means \rightarrow One\ Sample\ T\ test$

After clicking the **One-Sample T Test** option, you will be directed to the next screen for selecting variable for computing t-statistic (Figure 11.6). Select Khan_running variable from left panel and bring it to the right panel by clicking the arrow sign as **Test Variable**. In case of more number of variables, you may select them as well for computing t value for each variable. Use 20 as **Test value.**
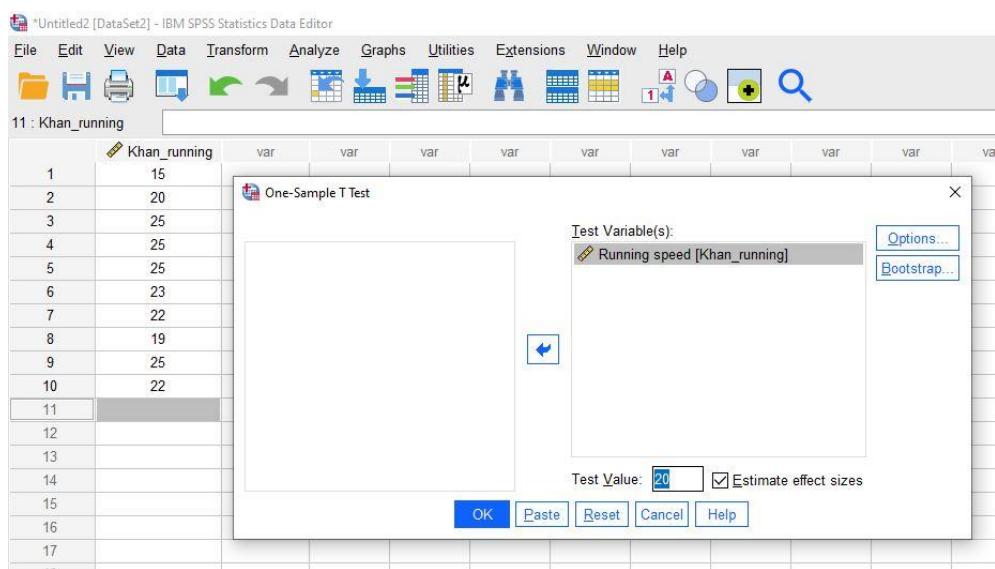
Figure 11.6: Data inputs (One sample T Test).

All the knowledge belongs to God.

Select the alpha level (**Confidence Interval Percentage**) from the **Options** tab (Figure 11.7) and press **Continue**:
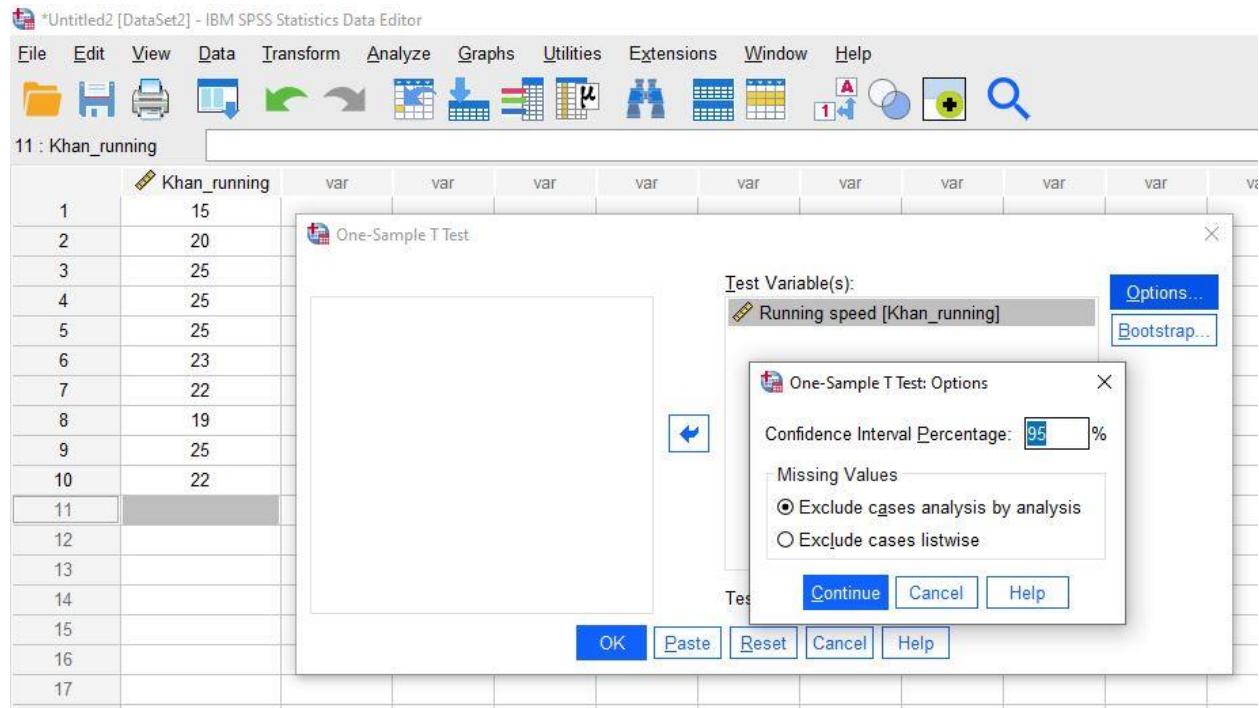
Figure 11.7: Options selection (One Sample T Test).

The output will be displayed as:
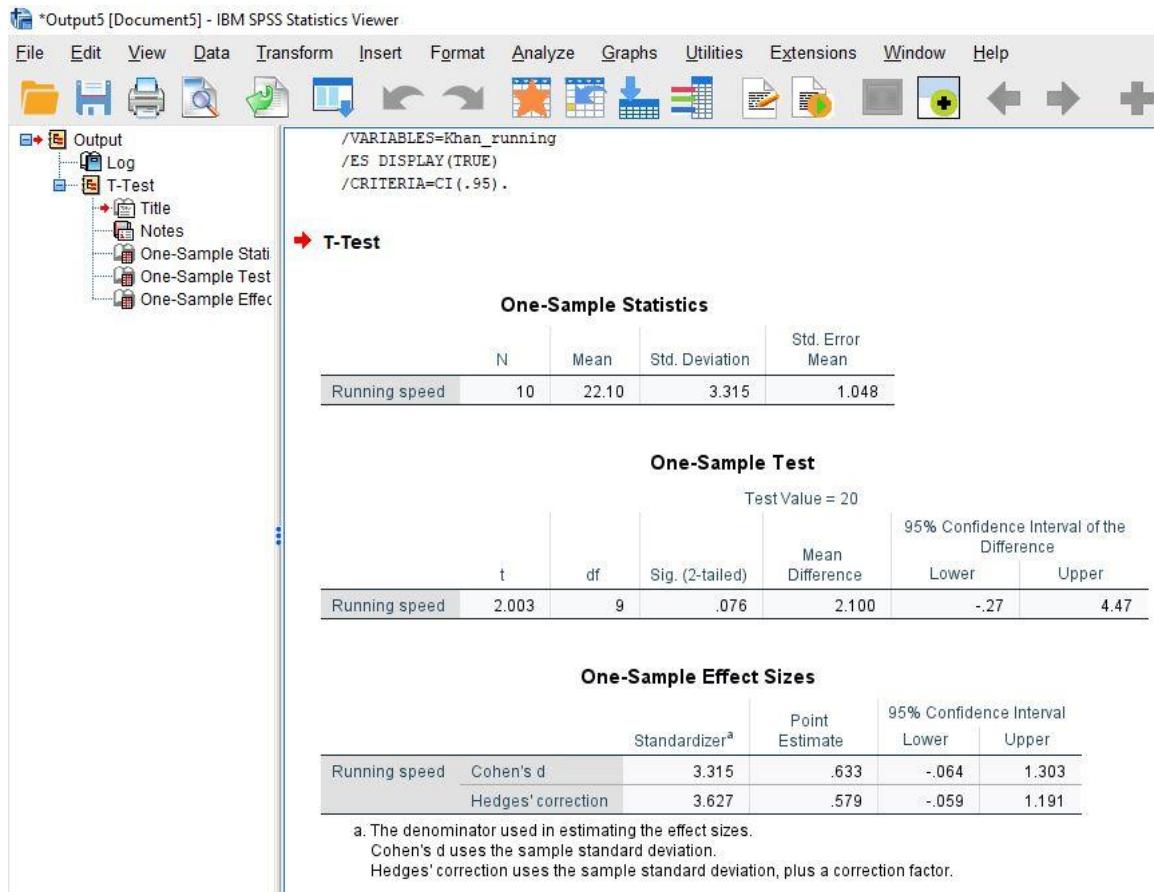
Figure 11.8: output (One Sample T Test).

Note the **Sig** value. It is (0.07) greater than P assumed (0.05), as the sig value is close to one, H_1 will be rejected (Khan's average running speed is 20 km per hour).

11.4: INDEPENDENT SAMPLES T-TEST

It is used to compare two independent or unrelated groups (**between-groups design**) on an approximately normal dependent variable. Samples are independent if there is no relationship between them (generally meaning each participant provides data for only one sample). In other words, an **Independent-samples t-test** is used to test whether the difference between two population means is significant.

All t-tests are usually called **Student's t-tests**. But strictly speaking, this name should be used only if the variances of the two populations are also assumed to be equal. In case the assumption of equality of variances is violated, then the Welch's t-test is used (Post Hoc).

It is the same as **Means** test but the difference is, here we have hypothesis testing among two means from different groups (independent). But we always assume that the variances of the two groups are equal (null hypothesis).

In two-tailed test if the null hypothesis is rejected, it may be concluded that the group means differ significantly and one cannot interpret as to which group mean is higher.

11.4.1: One-tailed and two-tailed

In one-tailed test, an experimenter is interested to verify whether the population mean is larger than or smaller than a given value, whereas in a two-tailed test, it is required

to know whether the population mean differs from the given value. Here, it is not of much interest to know the direction of difference. Simply put, *a statistical model that tests a directional hypothesis is called a* **one-tailed test**, *whereas one testing a non-directional hypothesis is known as a* **two-tailed test**.

A two-tailed hypothesis attempts to determine whether any difference (either positive or negative) exists. Thus, you have an opportunity to make a Type I error on either of the two tails of the normal distribution. A one-tailed test examines a difference in a specific direction. Thus, we can make a Type I error on only one side (tail) of the distribution.

11.4.2: Assumptions for T-test

While using the two-sample t-test the following assumptions are made:
- Population from which the samples have been drawn is normally distributed (normality).
- Variances of both the populations are equal (equal variances).
- Samples are independent to each other (independence).

11.4.3: Application

Consider an experiment in which the effect of two conditioning programs on fitness level needs to be compared. Two randomly selected group of subjects may be taken in the study. These two groups may be exposed to two different conditioning programs. Assuming that initial fitness level of both the groups is same, the null hypothesis of no difference in their final fitness scores may be tested by applying the two-sample t-test. In this case, both the samples are independent because the subjects in both the groups are different.

Similarly, consider a situation where a coach has developed two circuit training programs for his athletes and wish to know whether they differ in their effectiveness. Since he does not have an idea as to which program may be more effective, he would prefer to organize a two tailed test (two samples t-test).

11.4.4: SPSS Procedure

Let's say we have running average of Nasir and Khan (Table 11.3). We hypothesize that there is no difference among their means (null hypothesis). But we want to test it.

Table 11.3: Running speed of Khan and Nasir.

Nasir	120	110	99	100	140	133	130	140	144	138
Khan	100	110	98	95	88	88	88	85	90	95

The hypothesis we have:
$$H_1 = The\ average\ running\ speed\ of\ Khan\ is\ different\ from\ Nasir$$

The variables included in above example are:
- $Test\ variable = running\_speed$
- $Grouping\ Variable = coding\ for\ Nasir\ (as\ 1)\ and\ Khan\ (as\ 2).$

First of all, define the variable in **Variable View**. One of the conditions for using the two-sample t-test for independent groups is that the variance of the two groups must be equal. To do so, **Levene's F-test** shall be used to test the null hypothesis of equality of variances. If p value associated with the F-test is more than 0.05, the null hypothesis may be retained and this will ensure the homogeneity assumption required for using t-test. Put running data of both guys in single column into SPSS and using 1 and 2 as group coding in next column for **Grouping Variable** (1 for Nasir and 2 for Khan). After entering all the data in **Data View**, do the following steps:

All the knowledge belongs to God.

Analyze → Compare Means → Independent Samples T − Test
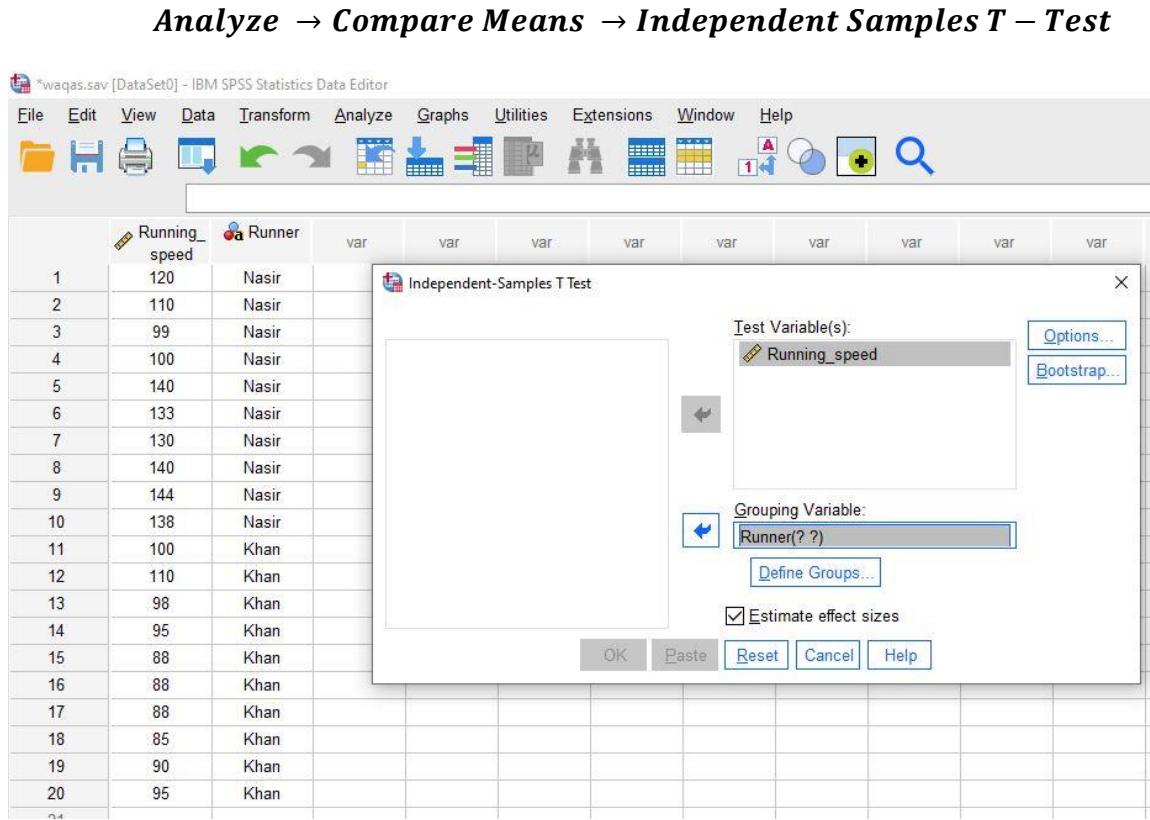
Figure 11.9: Variables inputs (Independent Samples T Test).

Our **Test variable** will be Running_speed and Runner as **Grouping Variable** (Figure 11.9). If you want to carry out t-tests on several dependent variables then you can select other dependent variables and transfer them to the variables list. However, there are good reasons why it is not a good idea to carry out lots of tests.

Grouping Variable can be defined using **Define Groups** tab (Figure 11.10). When your **Grouping Variable** has been selected the **Define Groups…** button will become active and you should click on it to activate the **Define Groups** dialog box. SPSS needs to know what numeric codes you assigned to your two groups, and there is a space for you to type the codes. In this example, we coded our Nasir's running as 1 and our khan's running as 2, and so these are the codes that we type.
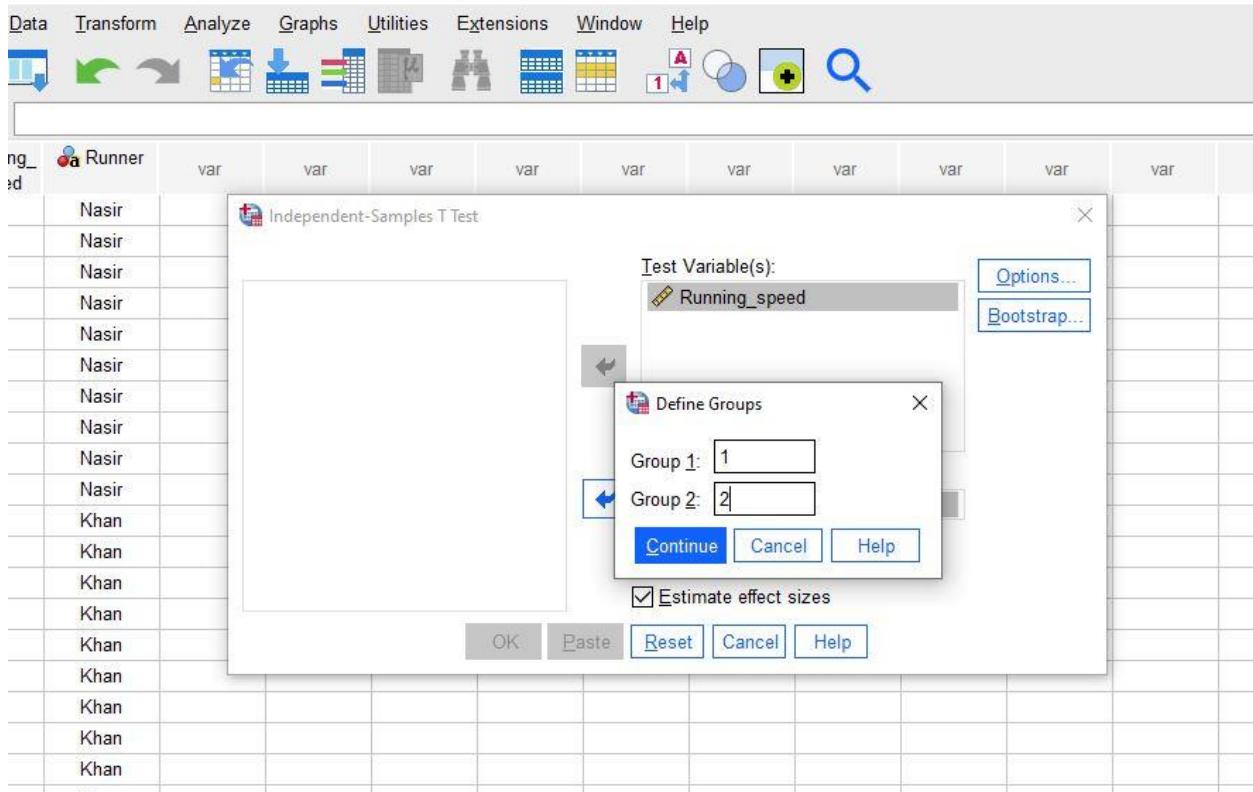
Figure 11.10: Define Group editing (Independent Samples T Test).

Alternatively, you can specify a **Cut point** in which case SPSS will assign all cases greater than or equal to that value to one group and all the values below the cut point to the second group. This facility is useful if you are testing different groups of participants based on something like a median split– you would simply type the median value in the box labelled Cut point. When you have defined the groups, click on to return to the main dialog box. If you click on then another dialog box appears that gives you the same options as for the dependent t-test.

Another feature in **Define Groups** is the **Cut point** which means that limit can be put on samples i.e., 90 minutes average to divide them into two groups (upper and lower). Set the **confidence level** from the **Options** tab and run the test by clicking **OK**. The output will be:
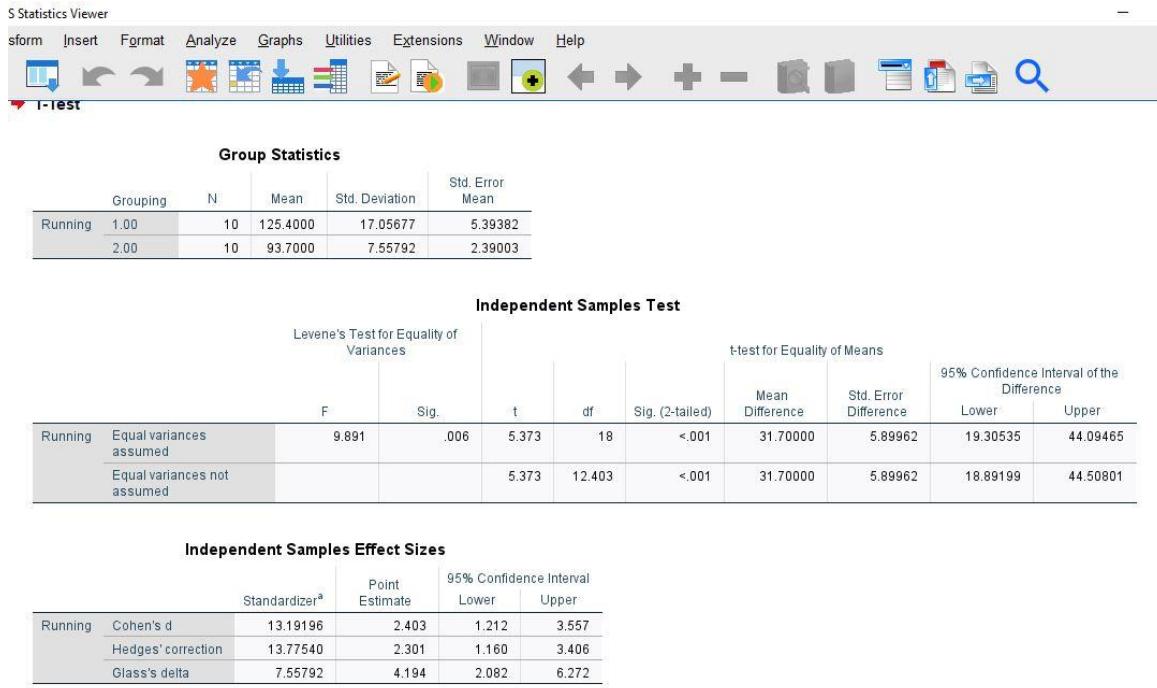
Figure 11.11: Independent Samples T Test output.

We have means and deviations etc. in first table (Figure 11.11). The second table is very important. The columns labeled **t**, **df**, and **Sig. (2-tailed)** provide the standard answer for the t test. They provide the value of t, the degrees of freedom (number of participants minus 2, in this case), and the **significance level** (often called p). Normally, we use the **Equal variances assumed** row.

Note that we have two rows under **Levene's test** for equality of variances: **equal variances assumed** and **equal variances nor assumed** (Figure 11.11). But we have **sig** value 0.006 for the first row (equal variances not assumed), assumptions of equal variances is violated. Levene's test for equality of variances assesses the assumption of homogeneity of variance. This test determines if the variation is similar or different between the groups. When Levene's test is not statistically significant, it means that we can continue with the regular independent-samples t test. In this situation, we can look at the independent-samples t-test table row that specifies that equal variances are assumed. When Levene's test is statistically significant, it means that there are differences in variation between the groups, so we have to make an adjustment to our independent-samples t test. In this situation, we can look at the independent-samples t-test table row that specifies that equal variances are not assumed.

We have **Sig** less than the P value under **Sig (2-tailed)** column (Figure 11.11), so we will reject the H_o and accept H_1: the averages of Khan and Nasir are different.

11.5: PAIRED SAMPLES T-TEST

It is used when the two scores being compared are paired or matched in some way (they are not independent of one another) or if the two scores are repeated measures. In other words, paired t-test is used to test a null hypothesis that the difference between two responses measured on the same subjects has a mean value of zero. Let us suppose we measure shooting accuracy of basketballers before and after a training program. If the

All the knowledge belongs to God.

training program is effective, we expect the shooting accuracy to improve for most of the basketballers after the training. Thus, to know the effectiveness of the program, a paired t-test is used. This paired t-test is also known as **repeated measures t-test** or **dependent t test**. *A repeated measures design is one in which measurements of the same variable are made on each subject on two or more different occasions.*

Remember that for comparison of means, both the means should be of same categories i.e. two means of head diameters can be compared but two different means (one for head diameters and one for weight) cannot be measured.

As we have collected data from one person/group or different groups give one data (1 mean from different groups or different means from one group). Thus, paired t-test is used in a situation where the subjects are same in pre- and post-testing. However, when the comparison is made between groups (of similar experimental units), it is called blocking. The paired difference experiment is an example of a randomized block experiment.

Pairing in an experimental design can serve to reduce bias, to increase precision, or both. Usually, the primary purpose of pairing is to increase precision. For example, in an institution, it has been observed that the research students are becoming lethargic due to lot of academic load and no compulsive physical activity. It is therefore, decided to launch a 40-min workout for them so that their muscular strength can be improved. Before launching the program it is decided to test the effectiveness of the program hence the workout may be given to 20 randomly chosen research students for 6 weeks. These subjects may be tested for their muscular strength by means of strength index before and after the 6-week workout program. In order to know the effectiveness of the workout, the paired t-test may be used.

11.5.1: Disadvantages of Paired T Test

The use of the paired comparisons test is not without its problems. If different subjects are used and randomly assigned to two treatments, considerable time and expense may be involved in our trying to match individuals on one or more relevant variables. A further price we pay for using paired comparisons is a loss of degrees of freedom. If we do not use paired observations, we have 2n-2 degrees of freedom available as compared to n-1 when we use the paired comparisons procedure.

In general, in deciding whether or not to use the paired comparisons procedure, one should be guided by the economics involved as well as by a consideration of the gains to be realized in terms of controlling extraneous variation.

11.5.2: SPSS Procedure

Ten PhD students participated in a running training program. Their speeds were measured before and after the program, which are shown in Table 11.4. Find whether the training program was effective at 0.05 significance level?

Table 11.4: Speed of students (Km/H) after 8 week training program.

Pre-prog	40	44	35	35	20	24	20	33	40	29
Post-prog	66	60	70	40	30	30	40	65	80	74

The hypothesis that needs to be tested is as following:

$H_1 = $ *The running speed after the management program is not the same as before.*

The variables included in above example are:
- $Variable\ 1 = \ Preprogram\ running\ speed$

All the knowledge belongs to God.

- *Variable* $2 = postprogram\ running\ speed$

For both of these variables, data shall be entered in two different columns unlike the way it was fed in **Two Samples T-Test** for unrelated groups.

Define the variables in Variable View first. Here the two variables preprogram_speed and postprogram_speed need to be defined in SPSS along with their properties. In **Data View**, click the following commands in sequence:

$$Analyze \rightarrow\ Compare\ Means\ \rightarrow Paired\ Samples\ T - Test$$

After clicking **Paired-Samples T Test**, the next screen will appear for variable selection. Once the dialog box is activated, you need to select pairs of variables to be analyzed. In this case we have only one pair (Preprog_speed vs. Postprogram_speed).

Select the variable Postprogram_speed and Preprogram_speed from left panel, and bring them to the right panel as **Variable 1** and **Variable 2** of pair 1.
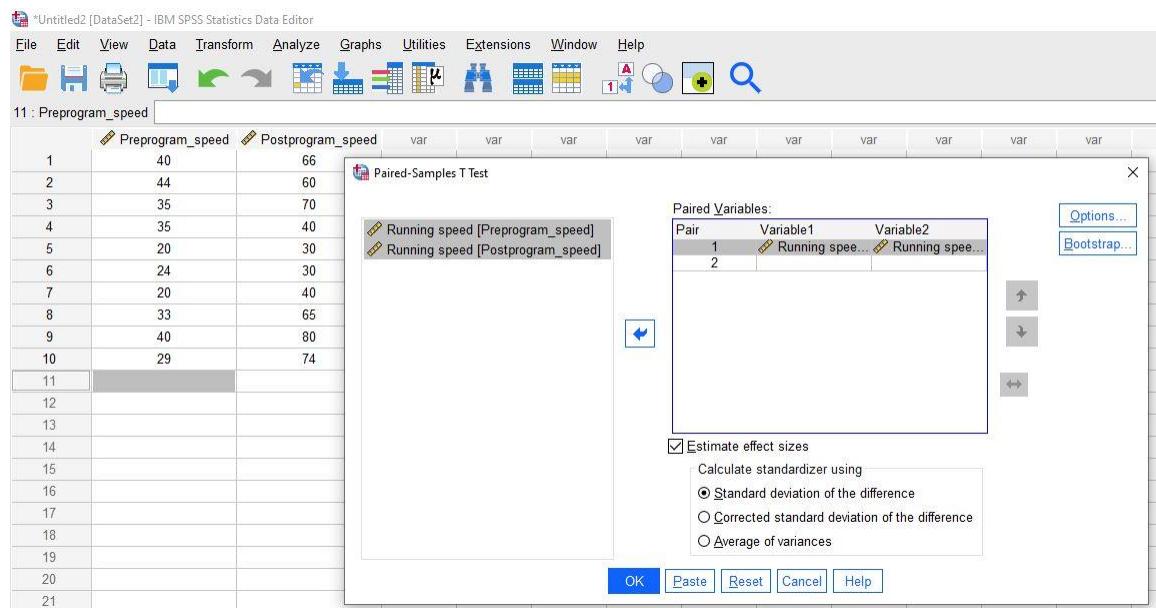
Figure 11.12: Paired Variables selection (Paired sample t-test).

Click on **Options** command, you will get the screen where confidence level is selected as 95% by default. You can change to any level. Click on **Continue**. Click on **OK** to get output. The output for the paired-samples t test consists of three components. The first part gives you basic descriptive statistics (**Paired Samples Statistics**) for the pair of variables as following:
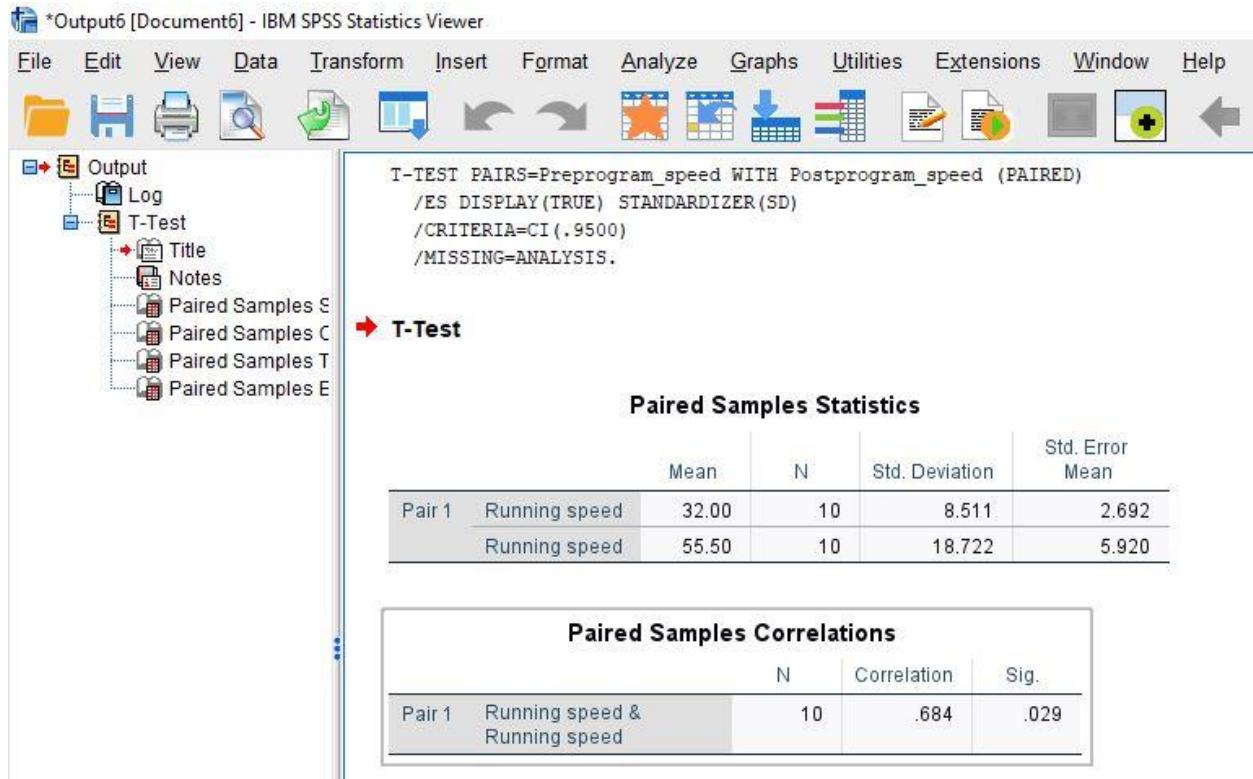
Figure 11.13: Paired sample test output (Statistics table).

The second part of the output (**Paired Samples Correlations**) is a Pearson correlation coefficient for the pair of variables (Figure 11.13).

Figure 11.14: Paired samples t-test output (Paired Samples Test table).

Within the third part of the output (labeled **Paired Samples Test**), the section called **Paired Differences** contains information about the differences between the two variables (Figure 11.14). The paired-samples t-test is essentially a single-sample t test calculated on the differences between the scores. The final three columns contain the value of t, the degrees of freedom, and the significance level.

All the knowledge belongs to God.

The t-statistic in table **(Paired Samples test)** is significant as its corresponding p value is <0.001, which is less than 0.05 (Figure 11.14). Thus, the null hypothesis of equality of mean speeds in post- and preprogram groups is rejected, and it may be concluded that the average speed of the students in post- and preprogram groups in the training program is not same.

11.6: ONE-WAY ANOVA

Also called single-factor ANOVA or CRD, is used when you have one independent variable with a few, often nominal, levels and one normally distributed dependent variable. The term "one-way" refers to the fact that there is one variable that defines the groups or treatments. Although three group means can be compared by using three t-tests.

For instance, one may like to compare the running time of three degrees: DVM, Poultry, and BBA students, or one may wish to compare the effects of breaststroke, butterfly stroke, and free stroke in swimming learning. Similarly, a study in which the effect of three different treatments (low, medium, and high intensity of circuit training program) on muscular strength is to be compared. The treatments have been randomly allocated to the subjects in such a manner that each treatment is received by an equal number of subjects. These all can be tested using One-way ANOVA.

11.6.1: Need of ANOVA

In general ANOVA procedures in SPSS produce what is called a Source Table in which each row represents a source of variability, and the columns represent what we know about the variability attributable to each source. In the simplest case total variability is split into that attributable to differences between groups and the variability attributable to differences within the groups.

Before explaining how ANOVA works, it is worth mentioning why we don't simply carry out several t-tests to compare all combinations of groups that have been tested. When we have three groups, we could use a t test to determine differences between the groups, but we would have to conduct three t tests (Group 1 compared to Group 2, Group 1 compared to Group 3, and Group 2 compared to Group 3). When we conduct multiple t tests, we infl ate the Type I error rate and increase our chance of drawing an inappropriate conclusion. ANOVA compensates for these multiple comparisons and gives us a single answer that tells us if any of the groups is different from any of the other groups.

Imagine a situation in which there were three experimental conditions, and we were interested in differences between these three groups. If we were to carry out t-tests on every pair of groups, then we would have to carry out three separate tests: one to compare groups 1 and 2, one to compare groups 1 and 3, and one to compare groups 2 and 3. If each of these t-tests uses a 0.05 level of significance then for each test the probability of falsely rejecting the null hypothesis (known as a Type I error) is only 5%. Therefore, the probability of no Type I errors is 0.95 (95%) for each test. If we assume that each test is independent (hence, we can multiply the probabilities) then the overall probability of no Type I errors is (0.95) 3 = 0.95 × 0.95 × 0.95 = 0.857), because the probability 3 of no Type I errors is 0.95 for each test and there are three tests. Given that the probability of no Type I errors is 0.857, then we can calculate the probability of making at least one Type I error by subtracting this number from 1 (remember that the maximum probability of any event occurring is 1). So, the probability of at least one Type I error is 1 - 0.857 = 0.143, or 14.3%. Therefore, across this group of tests, the probability of making a Type I error has increased from 5% to 14.3%, a value greater than the criterion accepted by social scientists.

All the knowledge belongs to God.

This error rate across statistical tests conducted on the same experimental data is known as the **familywise or experiment-wise error rate**.

11.6.2: SPSS Procedure (Equal Sample Size)

The data on anxiety obtained on students of DVM, Poultry, and Zoology students is shown in Table 11.5. Score was recorded from 0-100. The higher the score is higher the anxiety level. Find in which degree anxiety is higher. Discuss the findings at 5% level.

Table 11.5: Anxiety after different sports.

SN	DVM	Poultry	Zoology
1	22	25	25
2	21	20	30
3	21	19	28
4	23	20	20
5	22	16	26
6	23	18	29
7	21	21	33
8	24	16	33
9	22	17	27

In ANOVA, we try to compare between-group variability with that of within group variability.

$$H_1 = At\ least\ one\ of\ the\ group\ mean\ differs\ from\ others$$

The variables included in this example are:
- $Dependent\ Variable = anxiety$
- $Factor = Degree$

There are two variables in this example—namely, anxiety and degree that need to be defined along with their properties. Anxiety is a scale variable whereas Degree is a nominal variable: coded as 1, 2 and 3 for DVM, Poultry and Zoology respectively. Define these variables in **Variable View** as following:
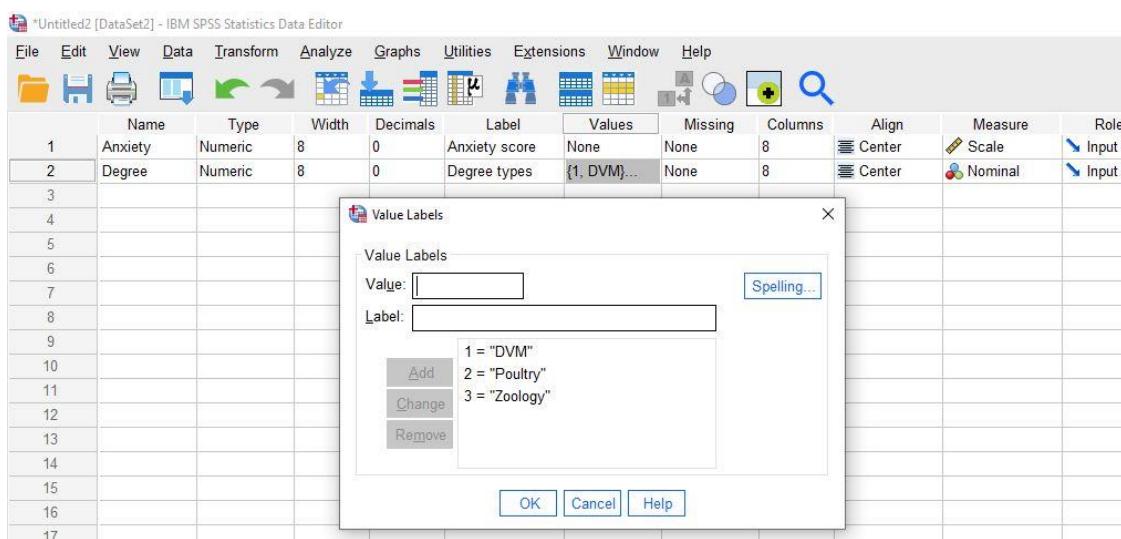
Figure 11.15: Defining variables in one-way ANOVA (Equal sample size).

All the knowledge belongs to God.

After entering all the data in the **Data View**, do the following steps:

$$Analyze \rightarrow Compare\ Means \rightarrow One\ Way\ ANOVA$$

Select the variables Anxiety and Degree from the left panel and bring them into the **Dependent list** section and **Factor** section in the right panel, respectively.
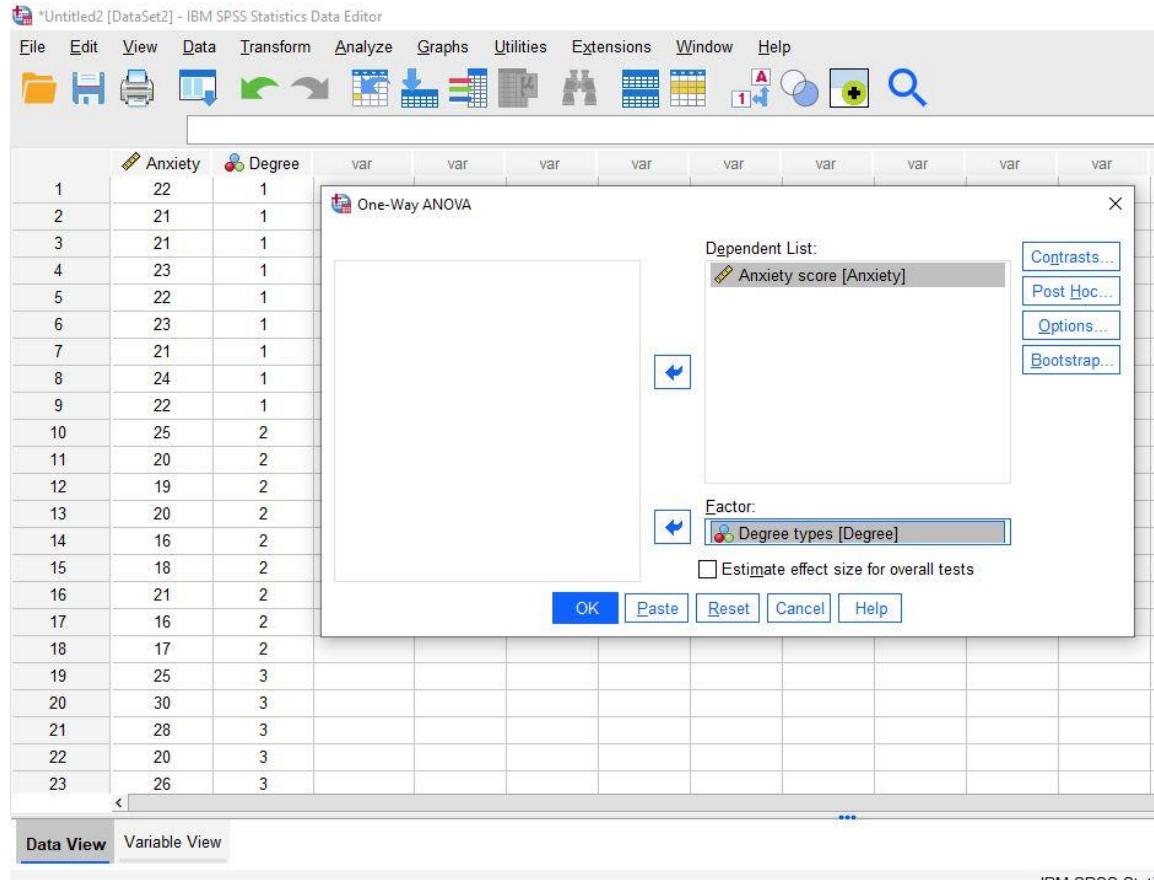
Figure 11.16: Variables input in One-way ANOVA (equal sample size).

Note that if there are more than 1 factors involve, it will be ANOVA problem even the levels are more than 1. Check test from the **Post Hoc** tab (Figure 11.17) to test which one is different from other as ANOVA does give this information.

All the knowledge belongs to God.
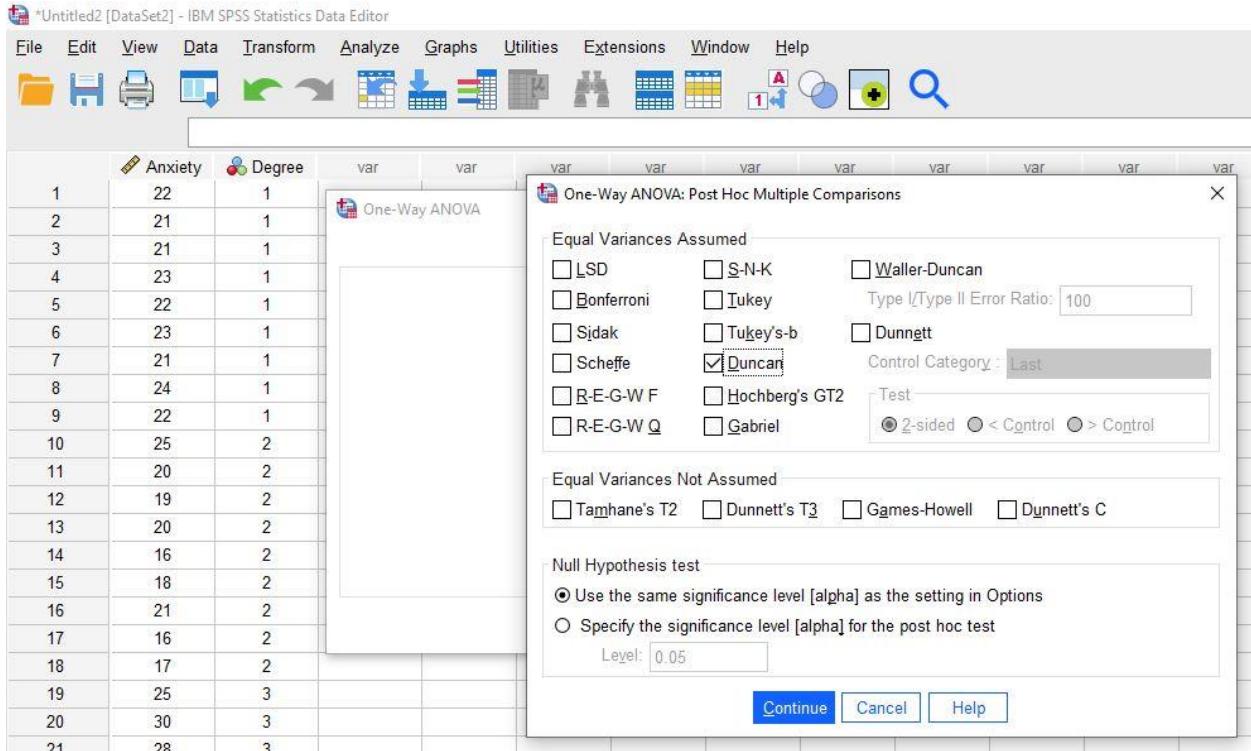
Figure 11.17: Post Hoc selection in one-way ANOVA (equal sample size).

Post hoc tests consist of pairwise comparisons that are designed to compare all different combinations of the treatment groups. So, *it is rather like taking every pair of groups and then per forming a t-test on each pair of groups*. Post hoc test needs to be applied for comparing means of groups. Post Hoc tests are appropriate when the researcher does not have a clear idea of which levels of the independent variable she/he wishes to compare or wants to compare all pairs of levels. Therefore, with Post Hoc tests all possible combinations will be compared.

Post-hoc tests are necessary in the event of a significant ANOVA. The ANOVA only indicates if any group is different from any other group. If it is significant, we need to determine which groups are different from which other groups. We could do t tests to determine that, but we would have the same problem as mentioned before with inflating the Type I error rate.

Use **confidence level** and other required parameters from the **Options** tab as following:
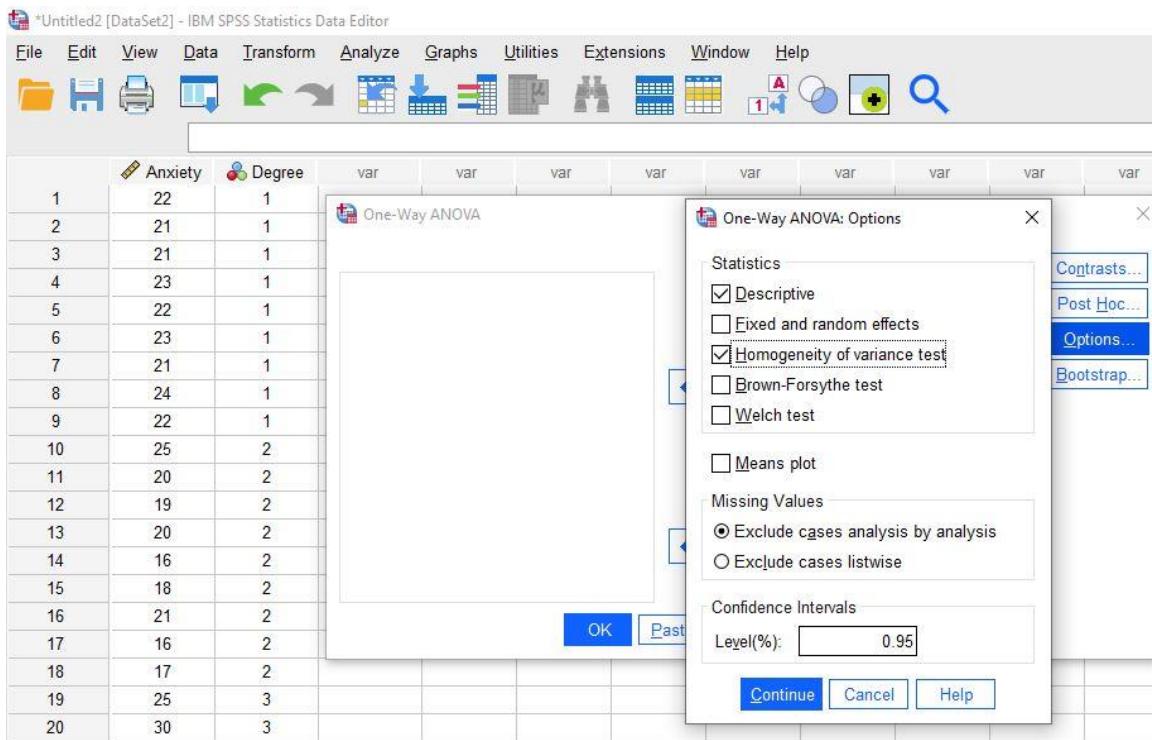
Figure 11.18: Options selection in one-way ANOVA (equal sample size).

Click **OK** to get the output:
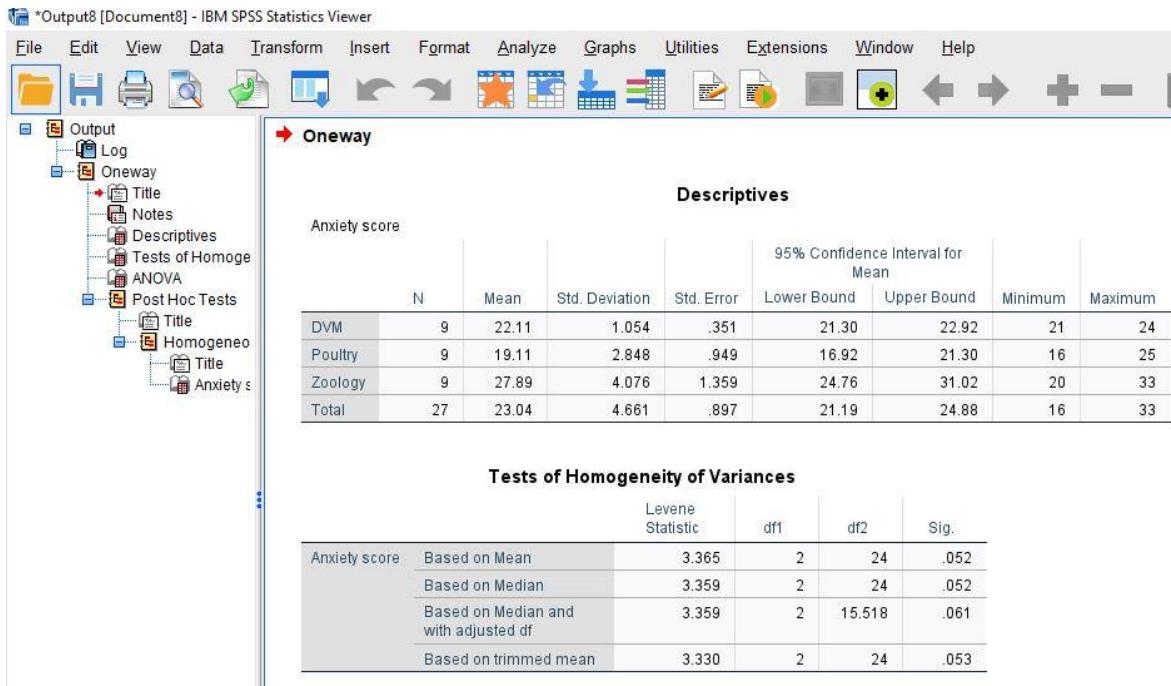
Figure 11.19: Descriptives and Homogeneity output in One-way ANOVA (equal sample size).

All the knowledge belongs to God.

The first two tables (**Descriptives** and **Test of Homogeneity of Variances**) are just as previously explained (Figure 11.19). Assumption of homogeneity is sustained in this case.

| *Output8 [Document8] - IBM SPSS Statistics Viewer | | | | | |

(SPSS output window showing ANOVA table)

Based on Median and with adjusted df		3.359	2	15.518	.001
Based on trimmed mean		3.330	2	24	.053

ANOVA

Anxiety score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	358.296	2	179.148	20.804	<.001
Within Groups	206.667	24	8.611		
Total	564.963	26			
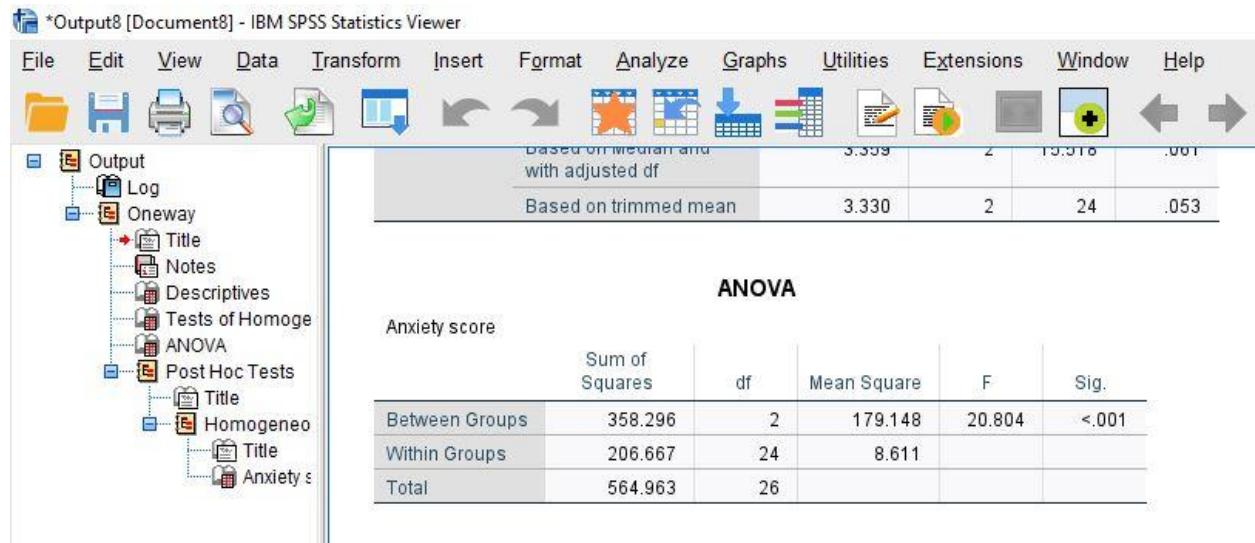
Figure 11.20: ANOVA table output in One-way ANOVA (equal sample size).

The next section of the output is the **ANOVA** source table (Figure 11.20). This is where the various components of the variance have been listed, along with their relative sizes. For a one-way ANOVA, there are two components to the variance: **Between Groups** (which represents the differences due to our independent variable) and **Within Groups** (which represents differences within each level of our independent variable). For our example, the Between Groups variance represents differences due to different degrees. The Within Groups variance represents individual differences among students.

While the terms 'between-groups' and 'within-groups' are not technical terms, they are useful in describing and understanding the ANOVA model (Figure 11.20). Computer software and other texts commonly refer to these sources of variability as **treatment (between groups)** and **error (within groups).** The row labelled within groups gives details of the unsystematic variation within the data (the variation due to natural individual differences in students and different reactions to degree types). The table tells us how much unsystematic variation exists (the residual sum of squares).

The **F** has two different degrees of freedom (df), one for **Between Groups** (in this case, 2 is the number of levels of our independent variable: $(3-1=2)$, and another for **Within Groups** (18 is the number of participants minus the number of levels of our independent variable: $(59-2=57)$.

Sig value of less than 0.001 indicates that the three means are different (accept H_1): p value (=0.001) is less than 0.05. Thus, the null hypothesis of no difference among means of the three groups may be rejected at 5% level.

All the knowledge belongs to God.

Post Hoc Tests

Homogeneous Subsets

Anxiety score

Duncan[a]

Degree types	N	Subset for alpha = 0.05		
		1	2	3
Poultry	9	19.11		
DVM	9		22.11	
Zoology	9			27.89
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 9.000.

Figure 11.21: Post-hoc output in One-way ANOVA (equal sample size).

The last table (Figure 11.21) in SPSS output shows the results of **Post-hoc (Duncan's test)**. This table presents us with every possible combination of levels of our independent variable.

This tests display subsets of groups that have the same means. Therefore, Duncan's test creates three subsets of groups with statistically similar means. The first subset contains the Poultry degree, second contains DVM degree and third contains Zoology degree (indicating that these three groups have the different means). These results demonstrate that all the degrees significantly vary from each other. Descriptive table indicates that the Zoology degree has highest anxiety among all degrees.

The tests provide a significance value for each subset and it's clear from these significance values that the groups in subsets have non-significant means (as indicated by values of Sig. that are greater than .05).

There are many post hoc tests available to compare the group means such as **least significance difference (LSD), Scheffe, Tukey, Bonferroni, Sidak, Duncan, etc. Tukey and Sidak are the most widely used** tests by the researchers. Duncan's test subdivided table into 2 subsets (columns). Which means that some means are significantly different from each other. If there were 3 subsets, it would have meant that all means differ and would be significant. If there were no significant difference, there would have been no table or no subsets. It may be seen that the mean anxiety of the Zoology is significantly higher (Figure 11.21) in comparison to that of DVM and Poultry. Remember that those groups which are under one subset (no one in this case) would not be significantly different from each other.

If the homogeneity of variance assumption is broken, then SPSS offers us two alternative versions of the F-ratio: the Brown–Forsythe F, and Welch's F.

11.6.2.1 Subset Arrangement in Post-hoc Test

There is a problem with ANOVA, it does not tell about which variable/level is different (better) from other. For example, look at the arrangement of Duncan under

Grouping, it has arranged the grouping as ascending order of means that is why it may be sometimes confusing when we have quantitative factor i.e. different protein levels (0%, 5%, 10% and 15%). However, if we have qualitative factor, the arrangement does not matter i.e., in present scenario of different 'degrees'. To avoid this error, we divide response into polynomial (first, second, third degrees polynomials; also known as **Polynomial Trends** or **Growth Curves**) because SPSS does not sort them as we want but as the order of means.

11.6.3: Polynomials

This tests for trends in the data and in its most basic form it looks for a linear trend (i.e. that the group means increase proportionately). However, there are more complex trends such as quadratic, cubic and quartic trends that can be examined. Figure shows examples of the types of trend that can exist in data sets.

Notice first that the linear trend is a straight line, but as the polynomials increase, they get more and more curved, indicating more rapid growth over time. Also, as polynomials increase, the change in the curve is quite dramatic (so dramatic that I adjusted the scale of the graph to fit all three curves on the same diagram). This observation highlights the fact that any growth curve higher than a quadratic (or possibly cubic) trend is very unrealistic in real data. By fitting a growth model to the data we can see which trend best describes the growth of an outcome variable over time (although, no one will believe that a significant fifth-order polynomial is telling us anything meaningful about the real world!).

If time is our predictor variable, then a linear trend is tested by including this variable alone. A quadratic or second-order polynomial is tested by including a predictor that is $time^2$, a cubic or third-order polynomial is tested by including a predictor that is $time^3$ and so on. So any polynomial is tested by including a variable that is the predictor to the power of the order of polynomial that you want to test: for a fifth-order polynomial we need a predictor of $time^5$ and for an n-order polynomial we would have to include $time^n$ as a predictor.

11.6.3.1 Constant Linear Response

For example, if we have different protein ratios (0, 5, 10 and 15%) to be checked on weight gain, and SPSS arranges them as 0, 5, 10 and 15% (after using Duncan test). We will say that the response is **Constraint Linear** or **First-Degree Polynomial**. First-order model has linear terms and takes the form: $Y = a + bx + cz$.

In other words, the linear trend should be familiar to you all by now and represents a simply proportionate change in the value of the dependent variable across ordered categories (the Figure 11.22 shows a positive linear trend but of course it could be negative).

All the knowledge belongs to God.
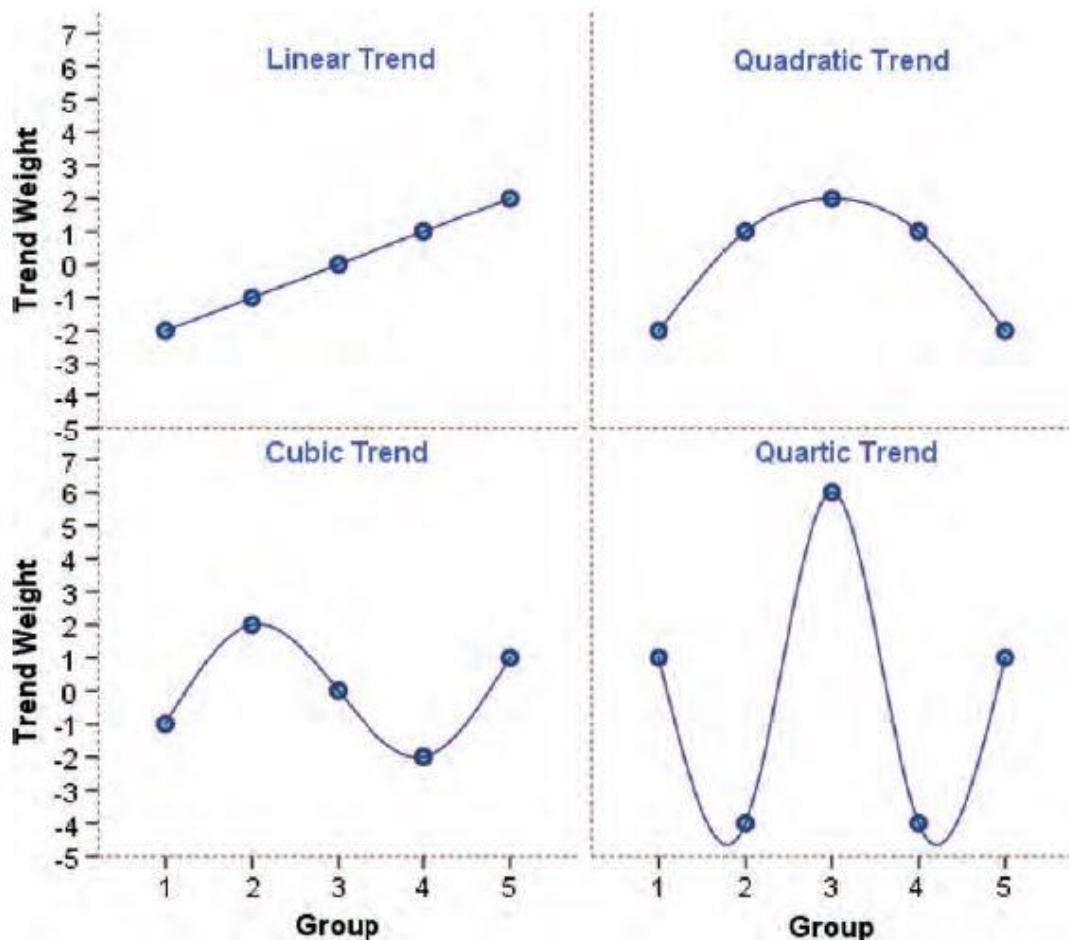
Figure 11.22: Polynomial trends.

11.6.3.2 Quadratic Response

A **Second-Degree Polynomial** or Quadratic Response (quadratic trend) is where there is one change in the direction of the line (e.g. the line is curved in one place) (Figure 11.22). An example of this might be a situation in which a drug enhances performance on a task at first but then as the dose increases the performance drops again. To find a quadratic trend you need at least three groups (because in the two-group situation there are not enough categories of the independent variable for the means of the dependent variable to change one way and then another). Suppose, if the response is 0, 5 15 and 10% arrangement (in subsets using Duncan test), it will be second degree polynomial. Second-order models include quadratic and cross product terms (product terms and in the first-order model, plus all quadratic terms like x^2 and all the terms xy).

11.6.3.3 Cubic Response

A Cubic Response or **Third Degree Polynomial** (cubic trend) is where there are two changes in the direction of the trend (Figure 11.22). So, for example, the mean of the dependent variable at first goes up across the first couple of categories of the independent variable, then across the succeeding categories the means go down, but then across the last few categories the means rise again. To have two changes in the direction of the mean you

All the knowledge belongs to God.

must have at least four categories of the independent variable. It is different altogether (**zig-zag**).

11.6.3.4 Quartic Trend

The final trend that you are likely to come across is the quartic trend (Figure 11.22), and this trend has three changes of direction (so you need at least five categories of the independent variable).
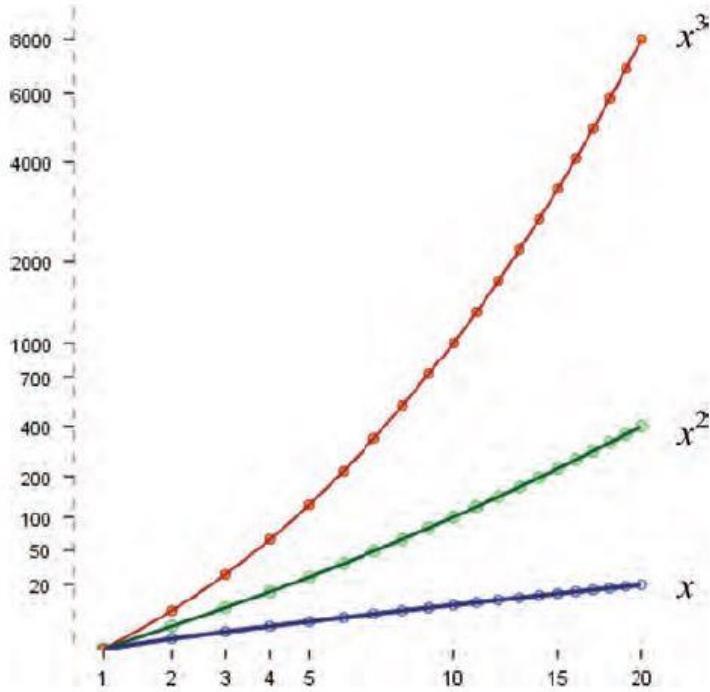
Figure 11.23: First-order (linear, blue), second order (quadratic, green) and third order (cubic, red) polynomial.

11.6.4: Linear Models

However, these above situations only arise when response is quantitative factor, for qualitative, it does not matter. For that purpose we divide a linear models into three types:
- Fixed models: all factors/treatments (even levels) are qualitative
- Random models: all factors are quantitative
- Mixed models: both types are present.

Classical ANOVA procedures were developed for fixed and random models.

11.6.4.1 Polynomial and Factor

Similarly, on the basis of number and nature of factors to be studied, we can divide design as following:
1) One factor: **only polynomial** (power) and no factorial
2) One factor: no polynomial (x^2 and y^2) **only factorial** (interaction).
3) Two factors: **polynomial factorial** (power + interaction), **it is even advanced than factorial design**. It gives best fitted model for study (maximum R^2 will give best result) using both power and interaction but through Design Expert.

All the knowledge belongs to God.

Factorial-treatment structures can be used for these kinds of experiments, but when treatment factors can be varied across a continuous range of values, other treatment designs may be more efficient. Response surface methods are designs and models for working with continuous treatments when finding optima or describing the response is the goal.

11.6.5: SPSS Procedure (Unequal Sample Size)

The average monthly university expense (in Pak Rupee) of different degrees, that is, DVM, BBA, Zoology and Poultry was obtained in a study, which is shown in Table 11.6. Apply test to find as to which category of students have the highest expense.

Table 11.6: Monthly expenses (unequal sample size).

DVM	BBA	Zoology	Poultry
16000	22000	10000	11000
15000	19500	9500	6000
18000	20000	9500	7600
18500	33000	9000	8500
16700	25000	10500	9900
13900	23000	12000	10500
20000	23000	15000	7000
19500		8000	7000
14000		11500	
15000		9000	

In this example, the null hypothesis that needs to be tested is:
$$H_1 = At\ least\ one\ group\ mean\ differs\ from\ others$$

The variables included are:
- $Dependent\ Variable = monthly\ expense$
- $Factor = degree\ (4\ levels)$
- $EUs = students$

There are two variables in this example—namely, Monthly Expenses and Degree Type that need to be defined along with their properties. Monthly Expenses is a 'scale' variable whereas Degree Type is a 'nominal' variable. Click on **Variable View** to define variables and their properties. Write short name of the variables as Expenses and Degree under the column heading **Name.** Under the column heading **Label,** full name of the two aforementioned variables may be defined as Monthly Expenses and Degree types.

Under **Values** column (Figure 11.24), code the degrees as 1 (DVM), 2 (BBA), 3 (Zoology) and 4 (Poultry) as following:

All the knowledge belongs to God.
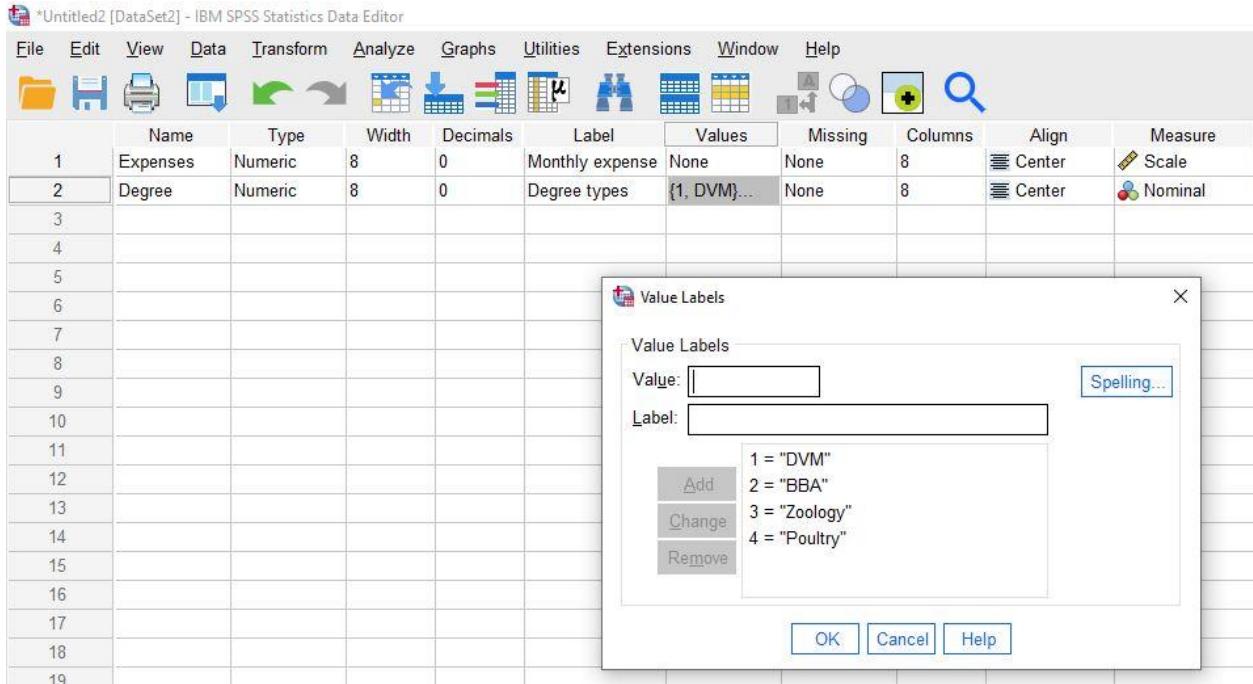
Figure 11.24: Defining variables in One-way ANOVA (unequal sample size).

In **Data View**, click the following commands in sequence:

$$Analyze \rightarrow \ Compare\ Means \ \rightarrow One\ Way\ ANOVA$$

After clicking **One-Way ANOVA** option, you will be taken to the next screen for selecting variables. Select variables Expenses and Degree from the left panel and bring them into the **Dependent list** and **Factor** sections in the right panel, respectively (Figure 11.25). Select **Post Hoc** and **Options** as explained previously.
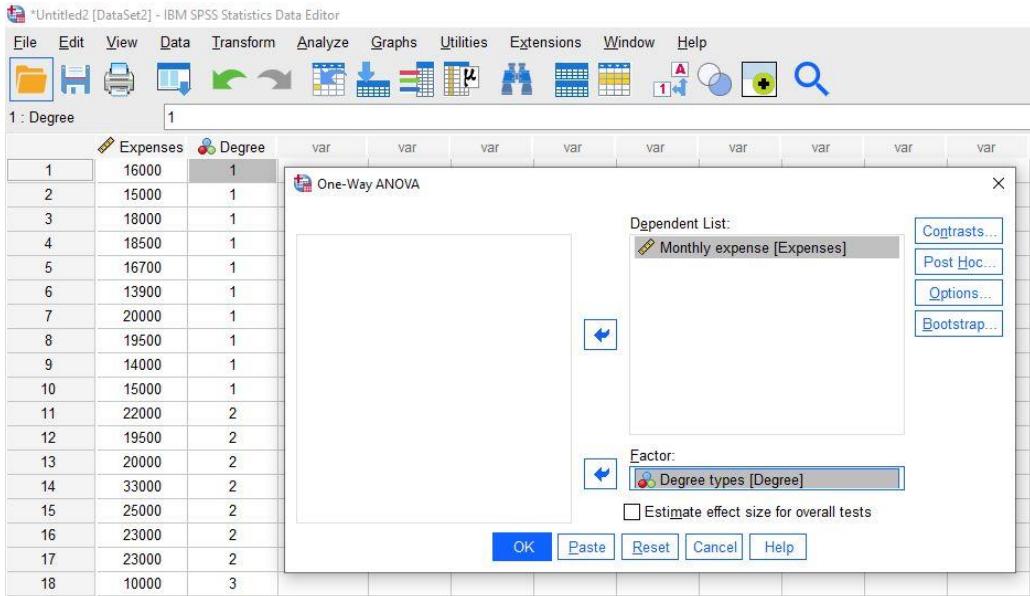
Figure 11.25: Variables input in One-way ANOVA (unequal sample size).

All the knowledge belongs to God.

If the null hypothesis is rejected, **post hoc test** shall be used for comparing group means from **Options** tab. Since the sample sizes are different, the **Scheffe test** shall be used for the **Post hoc** analysis. After selecting the variables, option needs to be defined for generating the outputs:
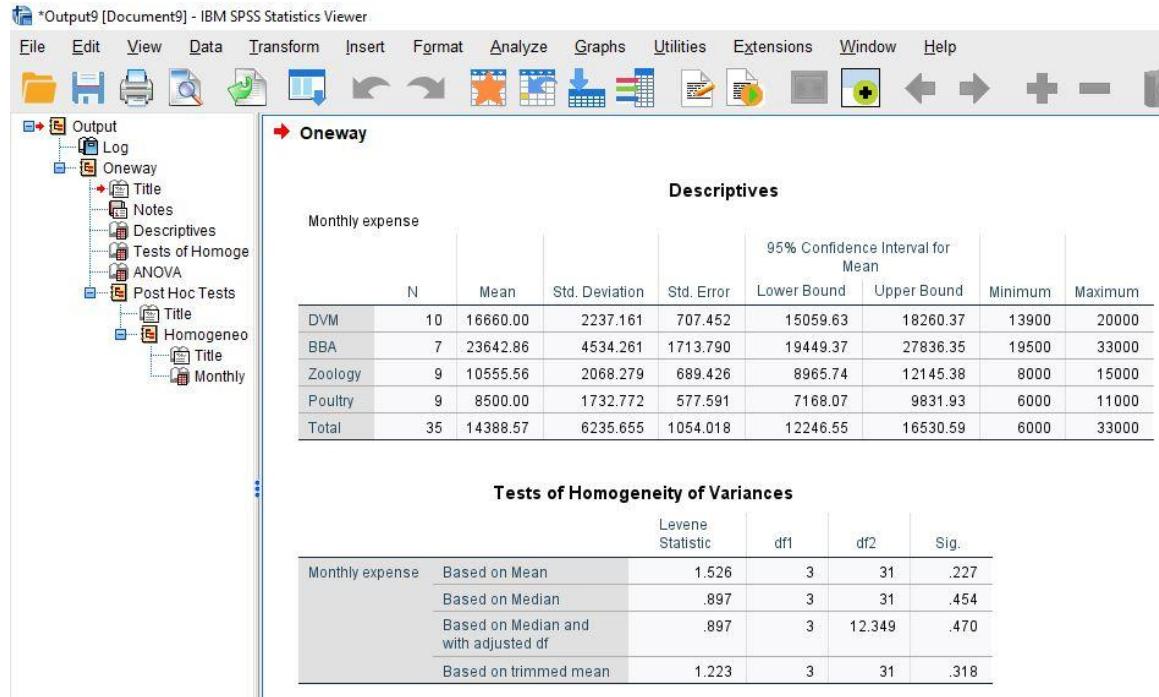
Figure 11.26: Descriptive and Homogeneity output in One way ANOVA (unequal sample size).

Descriptive table (Figure 11.26) provides highest and lowest means along with other parameters. BBA has the highest mean (monthly expense) among all the degrees (23642 Pak Rupees) and Poultry has lowest (8500).

Test of homogeneity is insignificant which means the equality of variance is sustained in this case.

All the knowledge belongs to God.
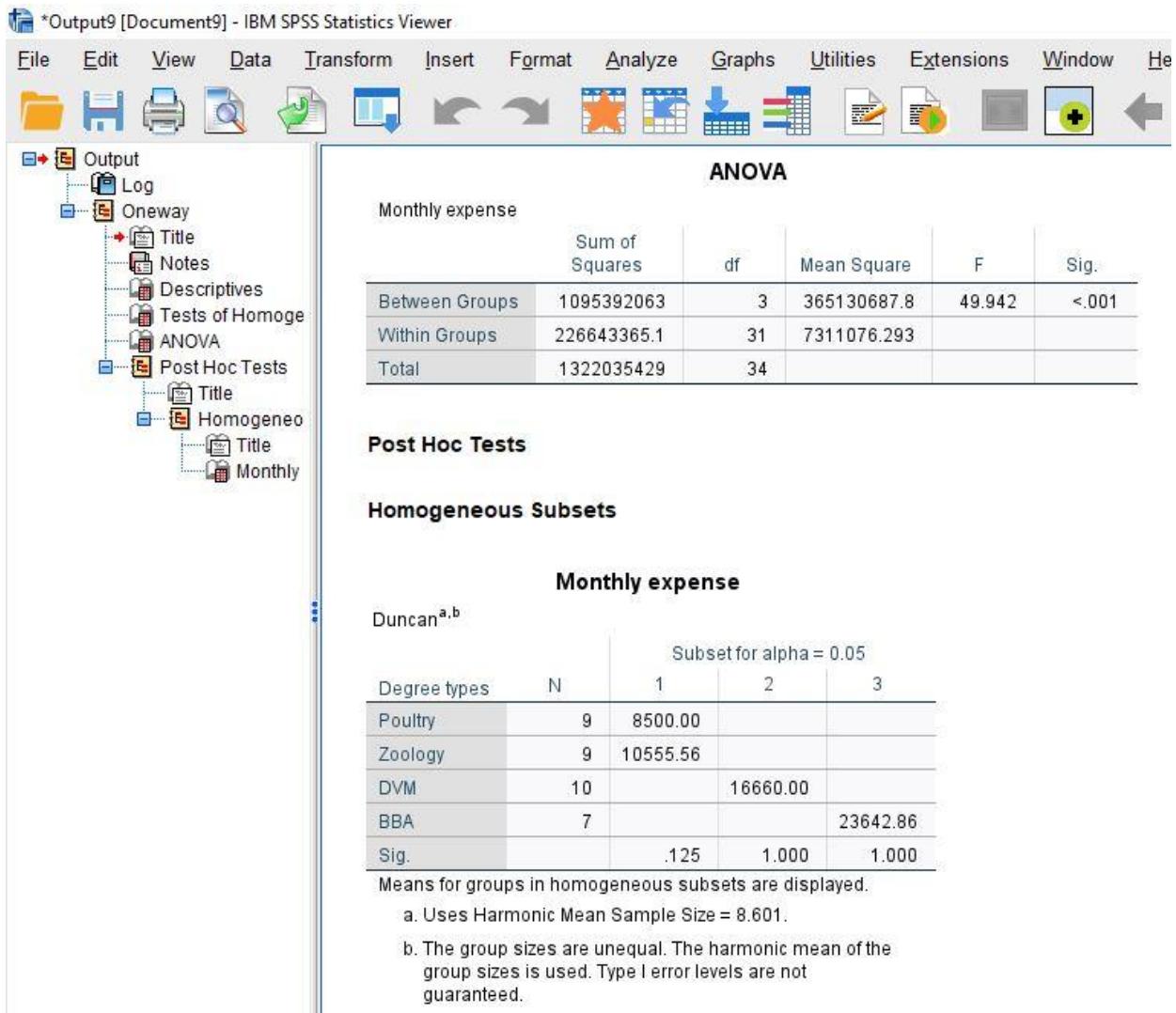
Figure 11.27: ANOVA and Post hoc output in One way ANOVA (unequal sample size).

P value (Figure 11.27) is less than 0.001 for **Between Groups** (degrees), which is less than 0.05. Thus, the null hypothesis of no difference among the means of the three groups, that is, DVM, BBA, Zoology and Poultry may be rejected at 5% level.

Post Hoc Tests makes the picture clearer. It is clear that the difference between degrees is significant. However, there is no difference between Poultry and Zoology in their monthly expenses. It can be concluded that the highest significant monthly expenses are with BBA.

CHAPTER 12: TWO-WAY ANOVA

12.1: INTRODUCTION

In one-way ANOVA, we have seen that the dependent variable is affected by the change in the different levels of an independent factor. On the other hand, in two-way ANOVA effect of two factors on dependent variable is investigated simultaneously**.** For instance, in studying the effect of different training intensities and weather conditions on muscular strength, the two factors that need to be investigated are training and weather. That is why it is also called **factorial ANOVA**. Remember that in SPSS, there is no RCBD, only two-way ANOVA is mentioned because it is an execution of design.

In above situation, one may conduct few simple experiments by using one-way ANOVA to investigate the following issues:

- Whether the impact of different training intensities on muscular strength is same in each weather condition
- Whether the impact of different weather conditions on muscular strength is same in each training intensity

Thus, if the impact of three training intensities and three weather conditions on muscular strength is to be studied, one has to organize six one-way ANOVA experiments. But, the impact of interaction, that is, joint effect of training and weather cannot be determined in such analysis. To overcome this problem and to utilize the experimental resources more economically, a two-way ANOVA experiment can be planned in this situation. In two-way ANOVA, there are two independent variables or factors (in this case training and weather) that affect the dependent variable (muscular strength). Further, it may be interesting to know as to which combination of treatment (training × weather) is the most effective proposition to enhance the muscular strength. A two-way ANOVA can be considered as an extension of one-way ANOVA. In such analysis, the effect of two independent factors on a dependent variable is studied; hence it is named as two-way ANOVA.

12.1.1: Hypotheses in Two-way ANOVA

In two-way ANOVA, the following three alternative hypotheses are tested:

a) The population means of all the levels of the first factor (main effect of first factor) are not equal. This is like the one-way ANOVA for the row factor.
b) The population means of all the levels of the second factor (main effect of second factor) are not equal. This is like the one-way ANOVA for the column factor.
c) There exists an interaction between the two factors. This is similar to performing a test for independence with contingency table.

Besides investigating main effects, two-way ANOVA facilitates the investigation of interaction effect between the two factors on dependent variable.

The number of treatment groups in the experiment is equal to the number of combinations of the levels of the two factors. For example, if the first factor has two levels and the second has three, then there will be $2 \times 3 = 6$ different treatment groups. In the example discussed earlier, let's assume that there are three different intensities of exercise and three different weather conditions. To see the impact of these two factors on muscular strength, there will be nine different treatment groups. Thus, 9 samples having the same size need to be identified so that these different combinations of treatments can be administered on them.

12.2: SPSS PROCEDURE (FACTORIAL)

Fifteen DVM and fifteen BBA students were randomly chosen for the study. In each category, the subjects were divided into three equal groups (5 students in each group). Three different types of trainings were randomly administered to these three groups of subjects for 4

All the knowledge belongs to God.

weeks. After 4 weeks of treatment, these subjects were given a fitness test (scored between 1 and 10), high score representing better performance. Test scores so recorded is shown in Table 12.1. Let us see how to apply test and interpret the findings:

Table 12.1: Fitness test scores of students.

		Training		
	Group	Intense	Medium	Light
Degree	DVM	10	8	5
		7	6	4
		9	8	7
		6	5	4
		8	6	5
	BBA	4	5	3
		4	4	3
		5	6	4
		2	7	2
		2	4	1

Here two main factors, namely, Degrees (factor A) and Trainings (factor B) as well as interaction between Degrees and Trainings (A × B) need to be studied. The following three hypotheses shall be tested:

$$H_1 = Change\ in\ Degree\ has\ an\ effect$$

$$H_1' = Change\ in\ Training\ has\ an\ effect$$

$$H_1'' = Interaction\ exists\ between\ Degree\ and\ Training.$$

The variables included here are:
- $Fixed\ Factor = degree\ and\ training$
- $Dependent\ variable\ (response) = fitness\ score$

To enter these data into the SPSS Data Editor, remember that levels of a between-group variable go in a single column. Applying this rule to these data, we need to create two different coding variables in the data editor to represent diet and sport. Once you have created the two coding variables, you can create a third variable in which to place the values of the dependent variable.

There are three variables in this example, namely, Fitness Score (response), Degree, and Training that need to be defined along with their properties (Figure 12.1). Fitness Score is a scale variable, whereas Degree and Training are nominal variables. Write short name of the Variables as Fitness_score, Degree, and Training under the column heading **Name.** Under the column heading **Label,** full name of these variables may be defined as Fitness test score, Degree types of the subject, and Training programs. Alternate names may also be chosen for describing the variables. Under the column heading **Measure,** select **Scale** option for the variable Fitness score and **Nominal** for Sport and Diet variables.

All the knowledge belongs to God.
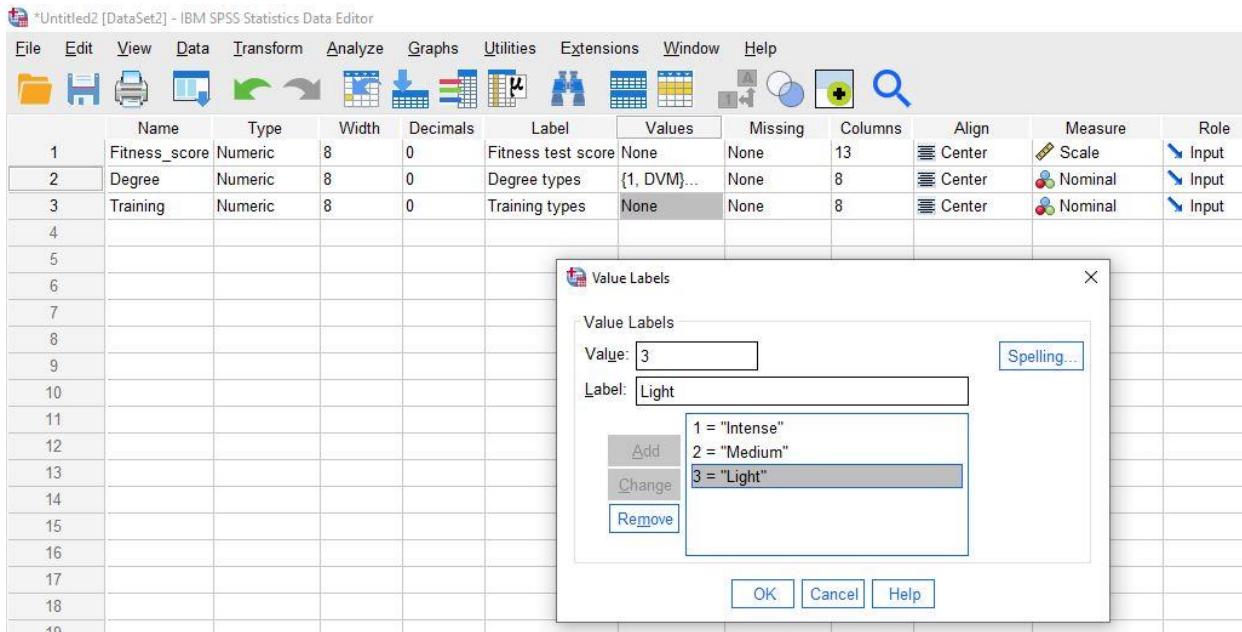
Figure 12.1: Defining variables (Two-way ANOVA).

Code the factor Training (1 for intense, 2 for medium and 3 for light) and Degree (1 for DVM and 2 for BBA) (Figure 12.1). Fitness scores in **Data View** are entered in the column of Fitness score. Under Degree first 15 scores are entered as 1 (denotes DVM students), next 15 scores are entered as 2 (denotes BBA). Under the column Training, first five scores are entered as 1 (denotes Intense training), next five scores as 2 (denotes Medium training), and next five scores as 3 (denotes Light training). In the **Data View**, click the following commands in sequence:

$$Analyze \rightarrow General\ Linear\ Model \rightarrow Univariate$$

After clicking the **Univariate** option, you will be taken to the next screen for selecting variables. Select the variable Fitness test score from left panel and bring it to the **Dependent Variable** section in the right panel. Similarly, select the Degree and Training variables from the left panel and bring them into the **Fixed Factor(s)** section in the right panel (Figure 12.2).

Remember that statisticians state that all common parametric statistics are relational. Thus, the full range of methods used to analyze one continuous dependent variable and one or more independent variables, either continuous or categorical, are mathematically similar. The model on which this is based is called the **general linear model (GLM).** This is an approach developed by **Nelder and Wedderburn** that essentially unifies linear and nonlinear models with both normal and non-normal responses.

All the knowledge belongs to God.
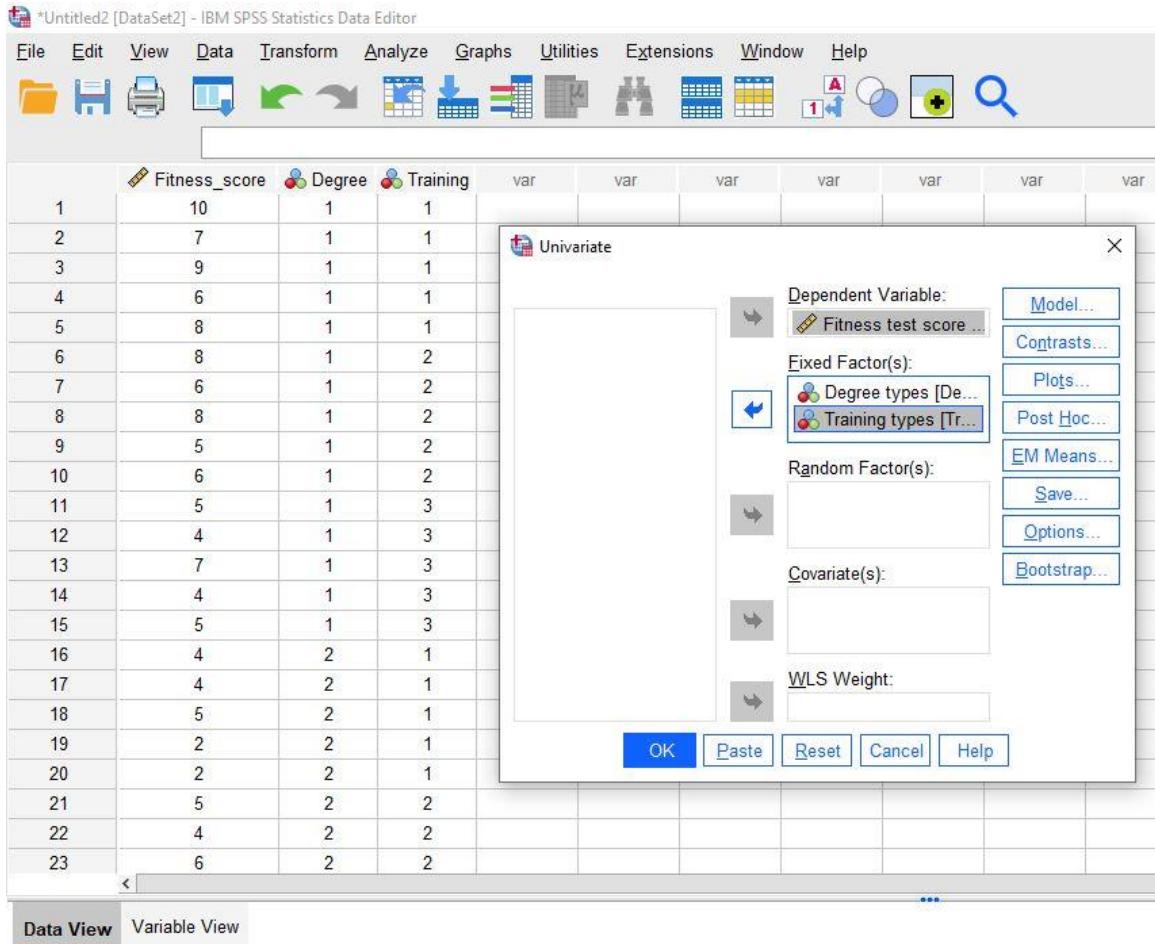
Figure 12.2: Variables input (Two-way ANOVA).

Click on **Post Hoc**, select the factors Training from the left panel, and bring it into the **Post Hoc Tests for** panel in the right side by using the arrow key (Figure 12.3). Check **Duncan** option. The variable Degree has only two levels and so we don't need to select post hoc tests for that variable (because any significant effects can reflect only the difference between DVM and BBA). However, there were three levels of the Training variable; hence we can conduct post hoc tests (although remember that normally you would conduct contrasts or post hoc tests, not both).
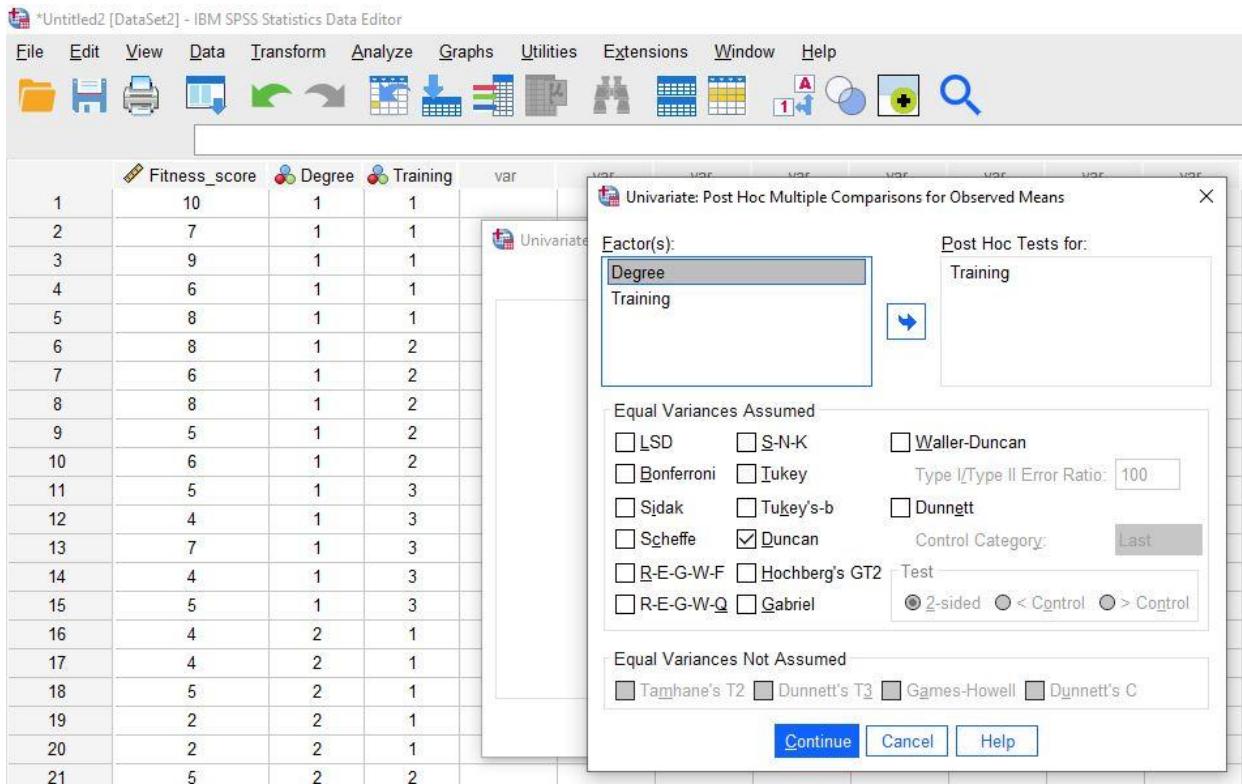
Figure 12.3: Post Hoc selection (Two-way ANOVA).

Click on **Continue**, now click on **Options** command and then check **Descriptive Statistics,** and **Homogeneity test** options (Figure 12.4). Ensure that the value of **significance level** is 0.05 in the box. Click on **Continue** to go back to the main screen.
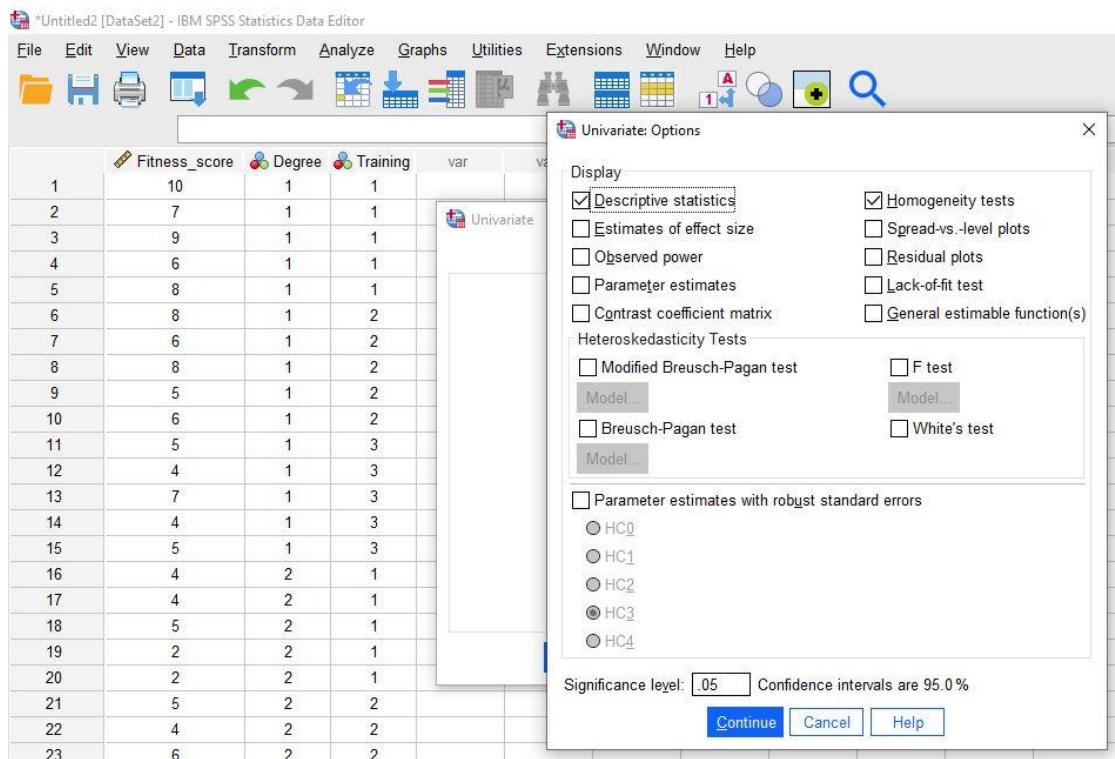
Figure 12.4: Selecting the Options (Two-way ANOVA).

Click on **Continue** and then on **OK** for generating outputs:
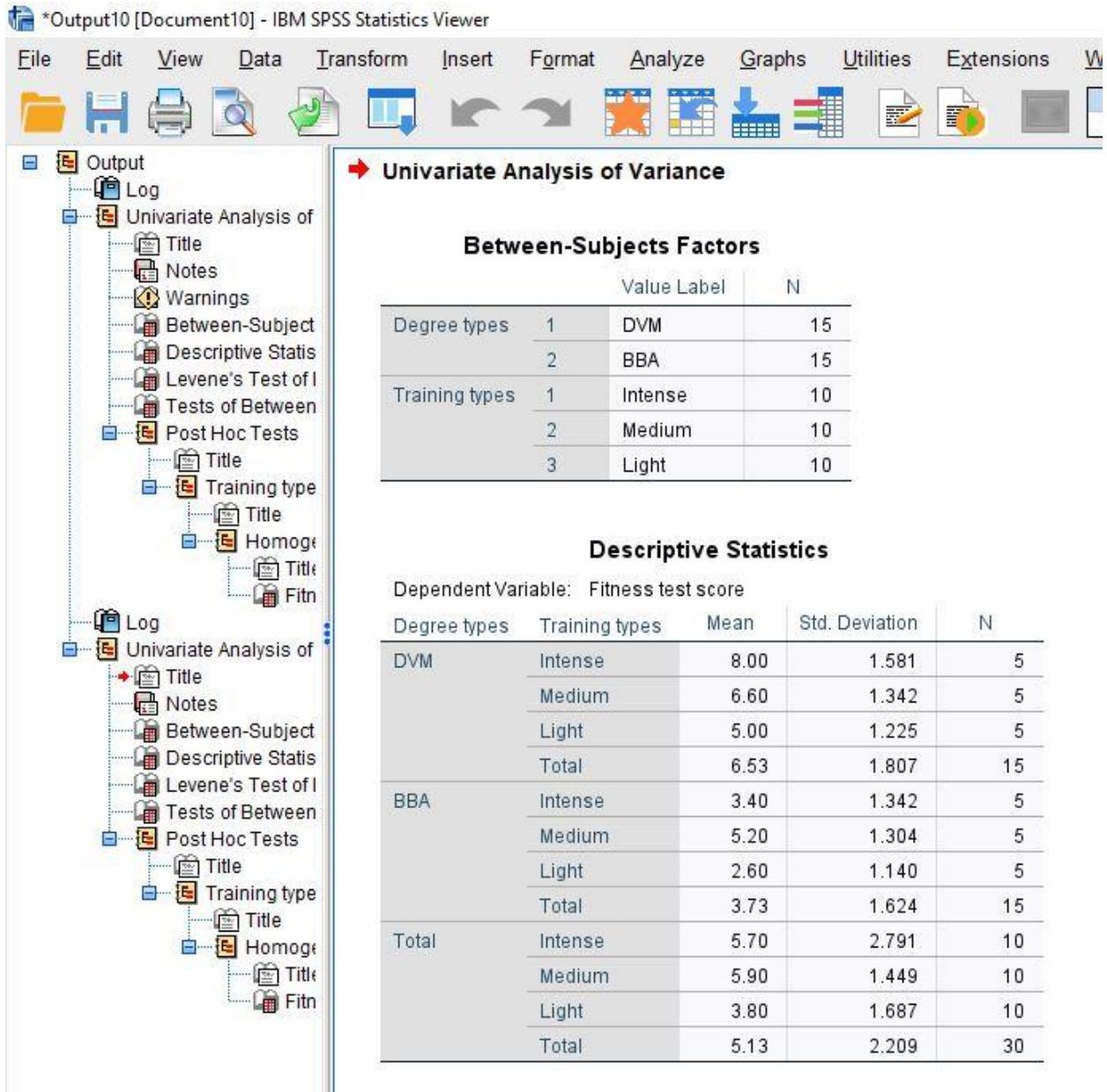
Figure 12.5: Descriptive Statistics output (Two-way ANOVA).

Between-Subject Factors table indicates the coding for each level of factor (Figure 12.5). The **Descriptives** table gives us the mean and standard deviations. Which indicates that the intense training of DVM has highest mean (8) among all levels.
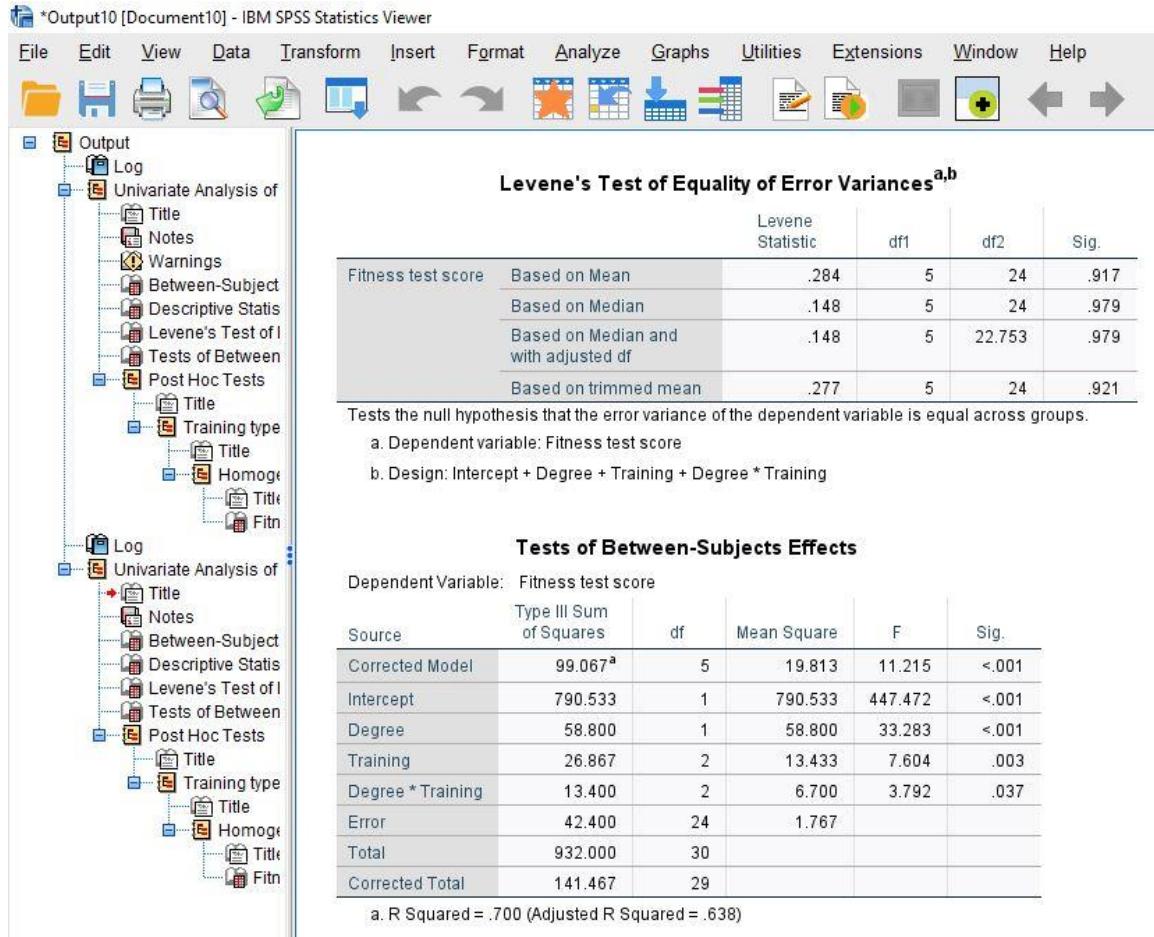
Figure 12.6: Levene's and Between-Subjects output (Two-way ANOVA).

In two-way ANOVA, one of the main assumptions is that the variance of the dependent scores across entire cell should be same. This can be tested by the Levene's test shown in Figure 12.6. This table reveals that the F value is not significant; hence the variability across all the cells is same. Thus, this assumption is satisfied.

Table (**Test of Between-Subjects Effects**) shows that the F values for Degree, Training, and Interaction are all significant because their associated values of p are less than 0.05 (Figure 12.6). Since interaction is significant, analyzing main effects of Degree and Training becomes meaningless.
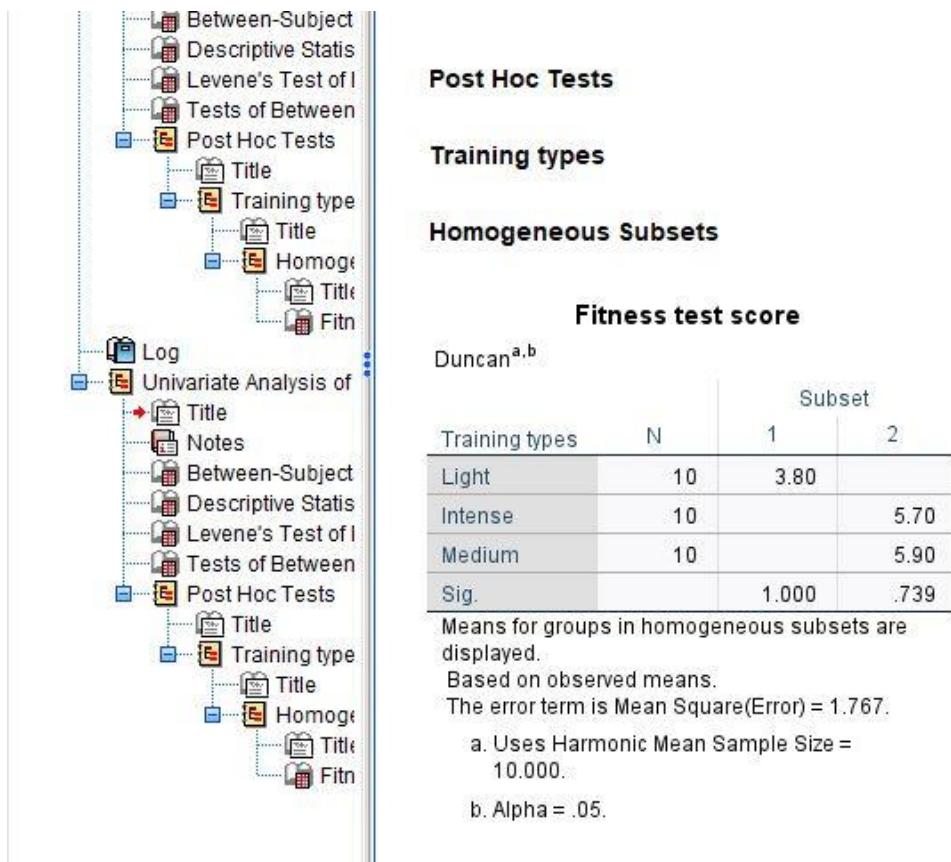
Figure 12.7: Post Hoc output (Two-way ANOVA).

It can be seen from Figure 12.6 that the effect of Training on fitness is significant; hence pair-wise comparison shall be done. Post hoc says that the Intense and Medium Training is more effective and that with Light is least effective for fitness irrespective of the Degree (Figure 12.7). However, Training with Intense and Medium Training is equally effective. In other words, training Intense and Medium are insignificant for each other but are significantly different from Light training.

12.3: MANOVA

MANOVA is multivariate analysis of variance. As such it is the extension of one-way ANOVA (univariate analysis). Multivariate tests are those that involve more than one response variable. While it is possible to conduct several univariate tests (one for each dependent variable), this causes Type I error inflation. Multivariate tests look at all dependent variables at once, in much the same way that ANOVA looks at all levels of an independent variable at once.

The SPSS procedure is slightly different as we use following commands to exist:

Analyze → General Linear Model → Multivariate

The rest of procedure is the same as two-way ANOVA. The output of MANOVA as multivariate test table is however slightly different. Four different types of multivariate test results are given: **Pillai Trace; Wilks' Lambda; Hotelling's Trace; Roy's Largest Root**. The most widely used is **Wilks' Lambda**.

12.4: THREE-WAY ANOVA

Three-way ANOVA is just like the two-way ANOVA, but there are 3 factors and three coding of one response in it instead of two factors and two coding of one response. Similarly, four-way ANOVA can be analyzed in SPSS. All the procedure is the same as two-way ANOVA.

12.5: DECODING FACTOR METHOD

Post-hoc can only tell which interaction is best but cannot tell which level of factor (say diet) is significant? We can convert two-way ANOVA problem to one-way ANOVA by decoding factor and check which level (code) is best/significant. It is actually a shortcut to use one-way ANOVA instead of two-way ANOVA. Similarly, 3-way ANOVA can be used and converted to one-way.

12.5.1: SPSS Procedure

Suppose we want to check the effect of three different diets on body weight of male and female chickens (Table 12.2). Use de-coding factor method to check the best level.

Table 12.2: Weight of chickens under 3 diets.

Gender	Diet		
	A	B	C
Female	2200	2100	2250
	2150	2100	2200
Male	2300	2400	2450
	2450	2400	2550

We will first make a **key** for de-coding purpose and then assign each level a code. As the above example has $3 \times 2 = 6$ levels, 6 coding keys will be made (Table 12.3). Remember to use decoding factor method technique only to the problem when there are few levels of each factor. For large number of levels, layout will be too large to handle on field.

Table 12.3: Decoding factors key (diet and gender).

Diet	Gender	Interaction	Key
A	Female	$A + Female$	1
B	Female	$B + Female$	2
C	Female	$C + Female$	3
A	Male	$A + Male$	4
B	Male	$B + Male$	5
C	Male	$C + Male$	6

Now we have only one factor (**Key**) instead of two factors (diet and gender) with 6 levels.

Enter the data for weight, gender and diet into the SPSS module and decode according to the key. Remember to define these variables in **Variable View** first (Figure 12.8) and mention the coding in **Values** for each variable especially for Key variable as mentioned in Table 12.3.

All the knowledge belongs to God.
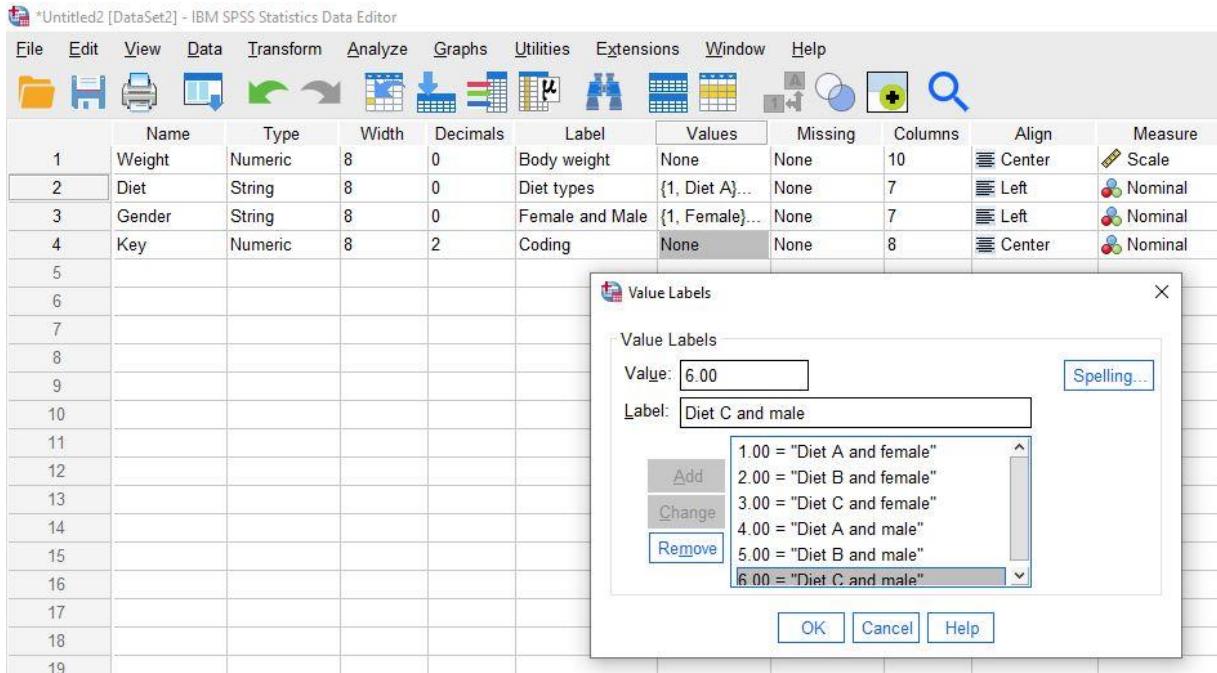
Figure 12.8: Defining variables (Decoding Factor).

Now use the same procedure as discussed for two-way ANOVA (univariate under GLM).
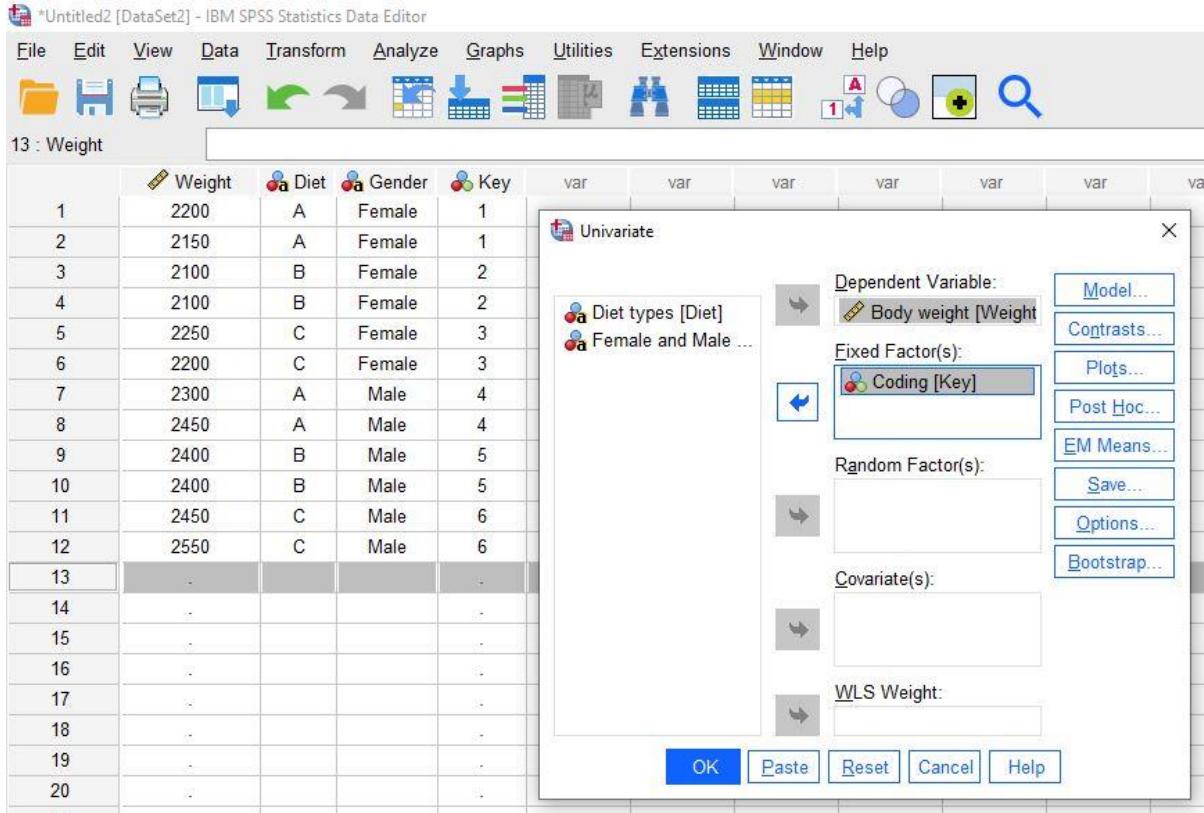
Figure 12.9: Variables input (Decoding Factor).

All the knowledge belongs to God.

Select the **Post-hoc** test **(Duncan)** for Key variable and then chose **Options**. Click **OK** to get output:
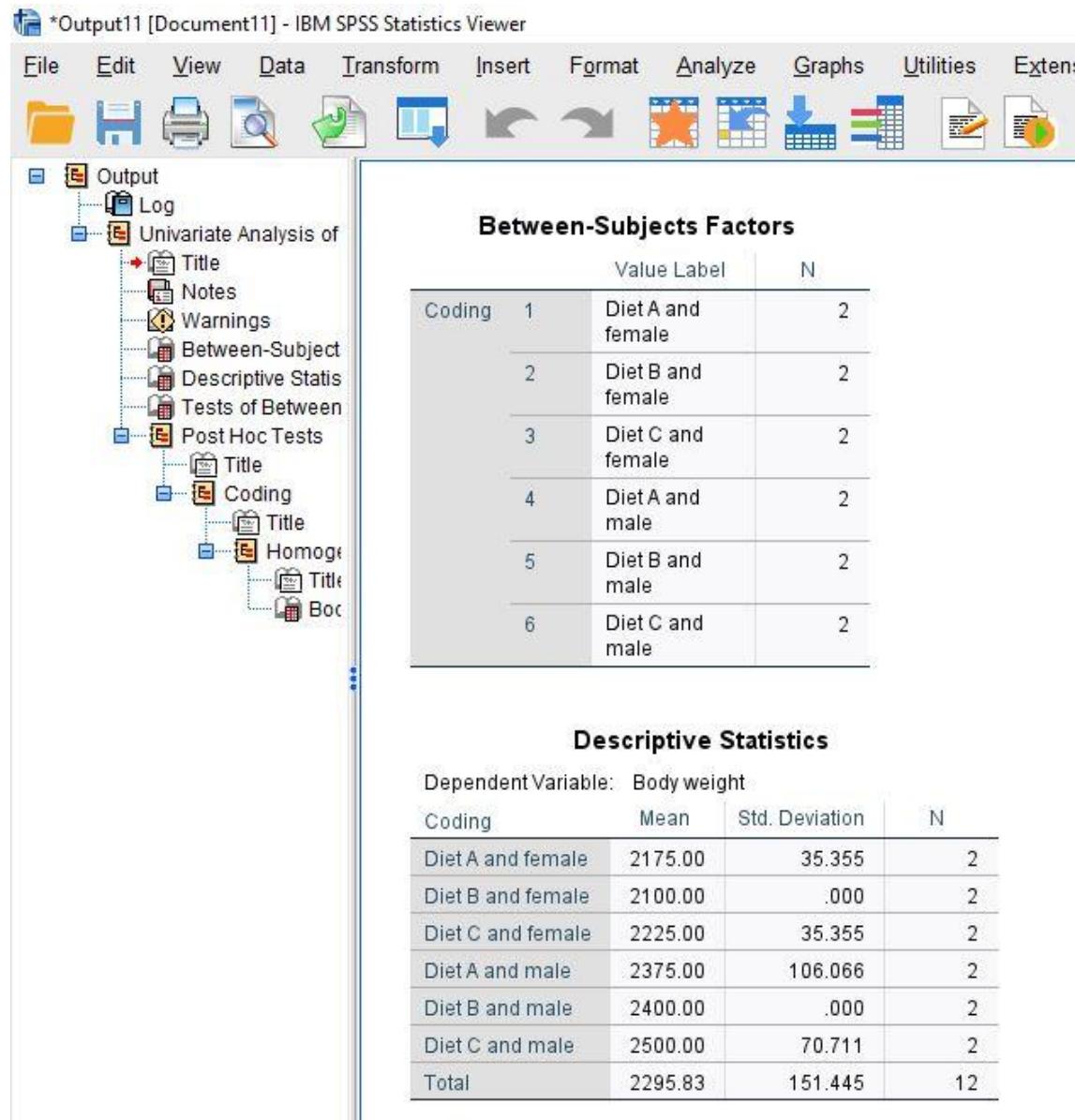
Figure 12.10: Descriptive Statistics output (Decoding Factor).

First table simply explains the coding done for **Key**. **Descriptives Statistics** table gives the idea about means (Figure 12.10).
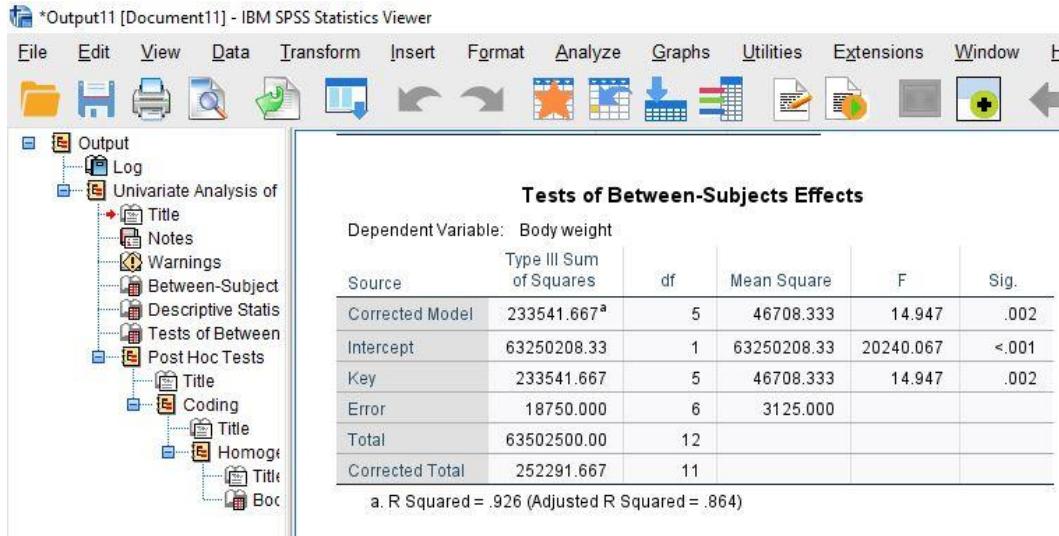
Figure 12.11: Tests of Between-Subjects Effect output (Decoding Factor).

Note that the **Key** is significant: ($P = 0.002$), which means that there is one level among our assigned levels which is significant (Figure 12.11). Now the Post-hoc will tell which one is the best.
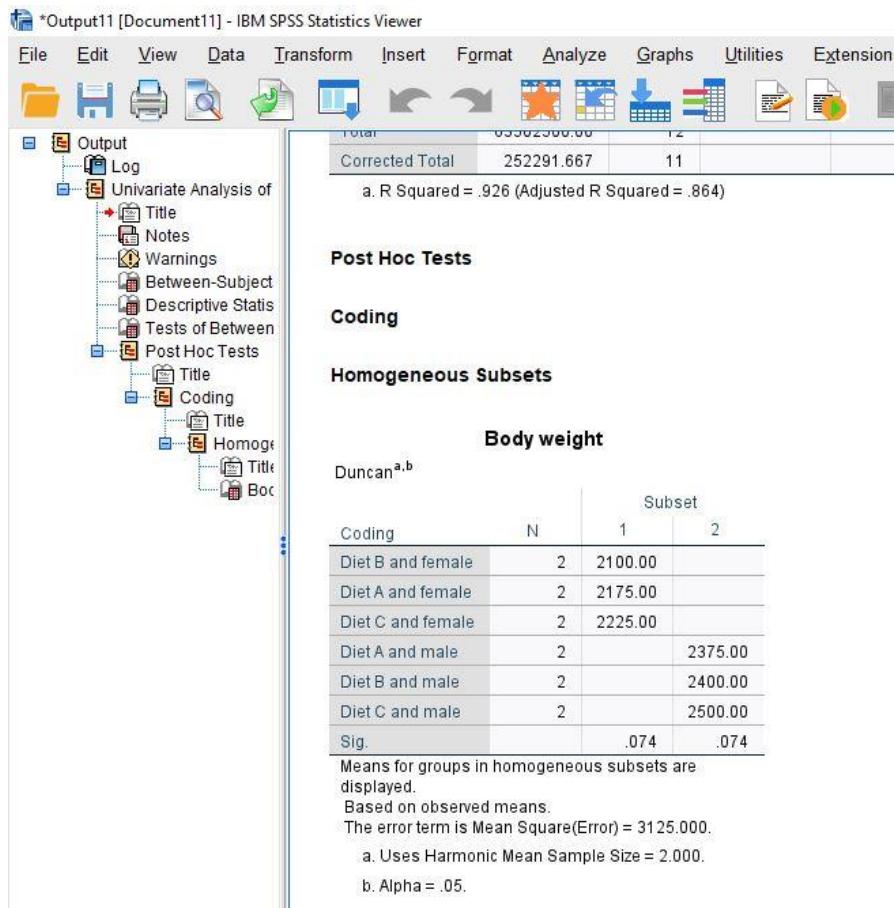
Figure 12.12: Post Hoc test output (Decoding Factor).

All the knowledge belongs to God.

Notice that Diet A, B and C all are insignificant for each other when combined with either female or male. However, the highest body weight is seen for males rather than female. Similarly, in subset 2 (Diet C and Male) is significant among all and has the best performance (highest weight=2500).

CHAPTER 13: REPEATED MEASURES ANOVA

13.1: INTRODUCTION

Many biological investigations involve repeated measurements made on the same individual under different conditions. These include studies of growth and development, studies of biological processes, and studies in which measurements are made before and after application of a certain treatment. Repeated-measures ANOVA extends the basic ANOVA procedure to a within-subjects independent variable. It functions like a paired-samples t-test when more than two levels are being compared.

Repeated measures is a term used when the same participants participate in all conditions of an experiment. *Repeated measure ANOVA (RMA) is used when at least one factor has response which is repeated after certain period of time (known as **within-subject factor**)* e.g., weight of chicken is taken after each week for 5 weeks after giving a certain diet. In repeated measures design, subjects serve their own control. For example, in order do away with the learning or fatigue effect (two treatments), sufficient time gap is maintained between any two treatments. Sometimes, a researcher may be interested to investigate the impact of training on the performance in different durations. For instance, one might be interested to know the effect of aerobic program on the VO_2 max after two, four, and six weeks, respectively.

Repeated measures design can be categorized as one way and two-way.

13.1.1: Within-Subjects versus Between-Subjects Designs

Within-subjects or repeated-measures designs are conceptually the opposite of between-groups designs. In within-subjects (sometimes called dependent) designs, each participant in the research receives or experiences all of the conditions or levels of the independent variable. These designs also include examples where the participants are matched by the experimenter or in some natural way (e.g., twins, husband and wife, or mother and child). When each participant is assessed more than once, these designs are also referred to as repeated-measures designs. The purpose of such experiment is to know the pattern of improvement among the subjects over a period of time. These within-group design are different from between-groups designs.

Between-groups designs are designs where each participant in the research is in one and only one condition or group. For example, there may be three groups (or levels or values) of the independent variable, treatment type. If the investigator wished to have 20 participants in each group, then 60 participants would be needed to carry out the research.

13.1.2: Pros and Cons of RMA

In experimental work in the social and behavioral sciences and some aspects of engineering the physical sciences, and business, the experimental units are frequently people. Because of differences in experience, training, or background, the differences in the responses of different people to the same treatment may be very large in some experimental situations. Unless it is controlled, this variability between people would become part of the experimental error, and in some cases, it would significantly inflate the error mean square, making it more difficult to detect real differences between treatments. In repeated measures design, one of the advantages is that variation due to subjects in different treatment groups is eliminated because each subject receives all the treatments.

Another advantage of these designs is that less number of subjects are required to perform the experiment. For example, that we have four treatments (in the usual sense) or four points in time on each of which we would like to have 10 measurements. If a different sample of subjects is used for each of the four treatments or points in time, 40 subjects

would be required. If we are able to take measurements on the same subject for each treatment or point in time—that is, if we can use a repeated measures design—only 10 subjects would be required.

A major potential problem to be on the alert for is what is known as the **carry-over effect**. When two or more treatments are being evaluated, the investigator should make sure that a subject's response to one treatment does not reflect a residual effect from previous treatments. This problem can frequently be solved by allowing a sufficient length of time between treatments.

Another possible problem is the **position effect**. A subject's response to a treatment experienced last in a sequence may be different from the response that would have occurred if the treatment had been first in the sequence. In certain studies, such as those involving physical participation on the part of the subjects, enthusiasm that is high at the beginning of the study may give way to boredom toward the end. A way around this problem is to randomize the sequence of treatments independently for each subject.

13.2: ONE-WAY RMA

In one-way RMA, an experimenter manipulates an independent variable to see the effect on some dependent variable where all the subjects participate in all the treatment conditions. Consider an experiment in which the effect of different interventions on the recovery pattern after the match is investigated among football players. The researcher may identify three different interventions (autogenic relaxation, aqua therapy, and yoga exercises) for a particular duration. In this design, same subjects are tested under each treatment condition to avoid the individual variation.

13.2.1: Assumptions in One-Way RMA

Using repeated measures design requires certain assumptions to be satisfied:
- The independent variable should be categorical and the dependent variable should be measured on interval or ratio scale.
- Observations obtained on the dependent variable must be independent from each other.
- The data on the dependent variable obtained on the subjects in each treatment condition must follow normal distribution.
- The sphericity should not exist among the data. Sphericity assumption is satisfied if correlations among the repeated measures on the dependent variable are all equal. If Mauchly's test is significant ($p<0.05$), sphericity assumption is violated.

The assumption of sphericity can be likened to the assumption of homogeneity of variance in between-group ANOVA. Sphericity is sometimes referred to as **circularity**. We assume that the relationship between pairs of experimental conditions is similar (i.e., *the level of dependence between experimental conditions is roughly equal known as sphericity*). It requires equal variances and covariances for each level of the within-subjects variable. Another way of thinking about sphericity is that, if new variables are created for each pair of within-subjects variable levels by subtracting each person's score for one level of the repeated-measures variable from that same person's score for the other level of the within subject variable, the variances for all of these new 'different' scores would be equal. For example, if you have three groups (Group A, Group B, and Group C) sphericity is assessing the following:

$$variance_{A-B} \approx variance_{A-C} \approx variance_{B-C}.$$

All the knowledge belongs to God.

13.2.2: Applications of One-way RMA

This design should be used in a situation where it is difficult to control the variation in the treatment groups due to individual variation. Some of these situations are as follows.

A researcher may wish to investigate the effect of different warming-up exercises on 400 meters event. A group of randomly selected athletes may be tested for their performance on 400 meter in each of the three treatment conditions; warm-up exercise with cold pack, hot pack, and mix of both on the abdomen. These within-group data may be compared by using the single-group repeated measures design.

The effect of conditioning program on fitness index may be studied on the subjects over a period of time. The purpose of such studies is to identify the time period in which significant improvement in the criterion variable occurs and also to know the time period after which the improvement stops increasing. The study may be planned in such a manner that each subject is tested for fitness index at 0 days and after 2, 4, 6, 8, and 10 weeks while undergoing the conditioning program.

In another situation, an investigator may study the effect of angle of release on shooting performance in basketball. A group of subjects may be tested for their performance on basketball shooting from a specific distance using three different angles: 45°, 50° and 55°.

An exercise scientist may like to investigate the effect of a low intensity exercise intervention on the cardio-respiratory endurance on overweight subjects. A random sample of subjects having weight 200 lb or more may be selected for the study on which the low intensity exercise program may be implemented. The performance of these subjects on cardio-respiratory endurance may be measured at 0 days and after 3, 6, 9, and 12 weeks of exercise intervention to identify the minimum duration for significant improvement. The investigator may further be interested to know the pattern of improvement in cardio-respiratory endurance during different time periods while undergoing the exercise intervention program.

13.2.3: SPSS Procedure

A statistician taught a subject, Experimental Design, with new technique to 8 randomly selected PhDs students, to see its impact on grades for assessing teaching efficiency. Repeated measures of marks were obtained on each subject at first, second, third, and fourth quiz which is shown in Table 13.1. Apply test to report its findings at the significance level 0.05.

Table 13.1: Data on marks (out of 10) obtained on the students.

First Quiz	2nd Quiz	3rd Quiz	4th Quiz
3	4	2	6
2	5	2	6
0	5	4	8
0	4	6	6
8	6	8	6
2	6	8	6
4	3	2	4
3	2	4	5

All the knowledge belongs to God.

Here it is required to test whether marks of the students differ in all the four quiz. To test this research hypothesis, the following alternative hypothesis shall be tested against null hypothesis that:

$$H_1 = New\ teahcing\ technique\ a\ progressive\ effect\ on\ marks$$
$$H_1 = Change\ in\ quiz\ has\ a\ progressive\ effect\ on\ marks$$

Variables included in this example are:
- $Within\ subjects\ factor = Teaching\ efficiency$
- $Response = Marks$

Define the variables in **Variable View** (Figure 13.1), because 'Teaching' is within-subject factor (repeated measures), so we will have to make four columns:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	First_quiz	Numeric	8	0	Marks	None	None	10	≡ Center	⬦ Scale	↘ Input
2	Second_quiz	Numeric	8	0	Marks	None	None	10	≡ Center	⬦ Scale	↘ Input
3	Third_quiz	Numeric	8	0	Marks	None	None	8	≡ Center	⬦ Scale	↘ Input
4	Fourth_quiz	Numeric	8	0	Marks	None	None	10	≡ Center	⬦ Scale	↘ Input
5											

Figure 13.1: Defining variables for within-subject factor (one-way RMA).

While being in the **Data View**, click the following commands in sequence:
$$Analyze\ \rightarrow\ General\ Linear\ Model\ \rightarrow\ Repeated\ Measures\ ANOVA$$

After clicking Repeated Measures command, the screen shall be obtained to define the variables. By default, the **Within-Subject Factor Name** is written as **factor 1**. Change this by Quiz because this is the independent (within-subjects) variable in this example (Figure 13.2). Write the number of **Levels** as 4, as there are four time periods in which the data has been obtained. Click **Add**. The variable names should be ordered according to when they occurred in time (i.e., the values of the independent variable that they represent).

In the **Measure Name** area type 'Marks'. Click **Add** (Figure 13.2). Please note that the name of the independent and dependent variables should start only from alphabet, and no gap should be there in between the two words in defining these names. If name contains two or more words, they must be joined by using the underscore.

All the knowledge belongs to God.
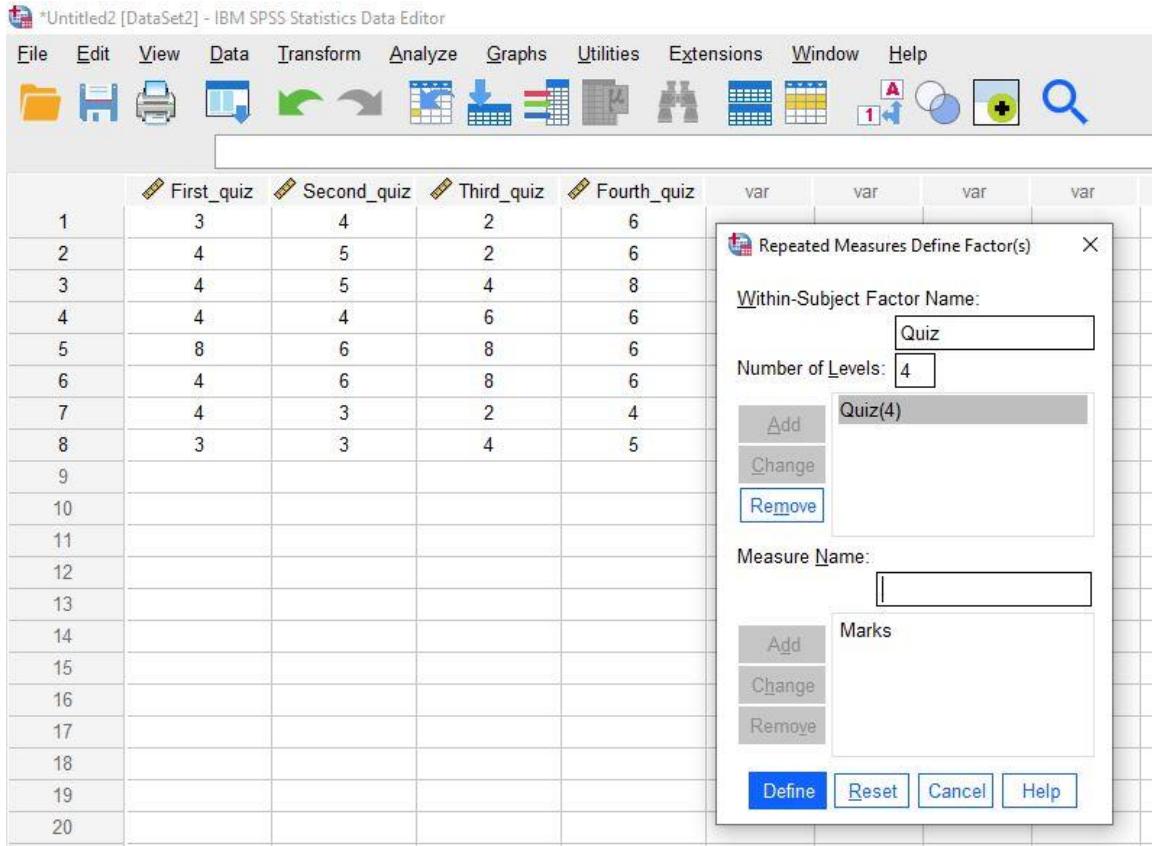
Figure 13.2: Variables input (One-way RMA).

Clicking **Define** in the screen will take you to the screen for selecting the within-subjects variables (Figure 13.3). The main dialog box has a space labelled **Within-Subjects Variables** that contains a list of four question marks followed by a number. These question marks are for the variables representing the four levels of the independent variable.
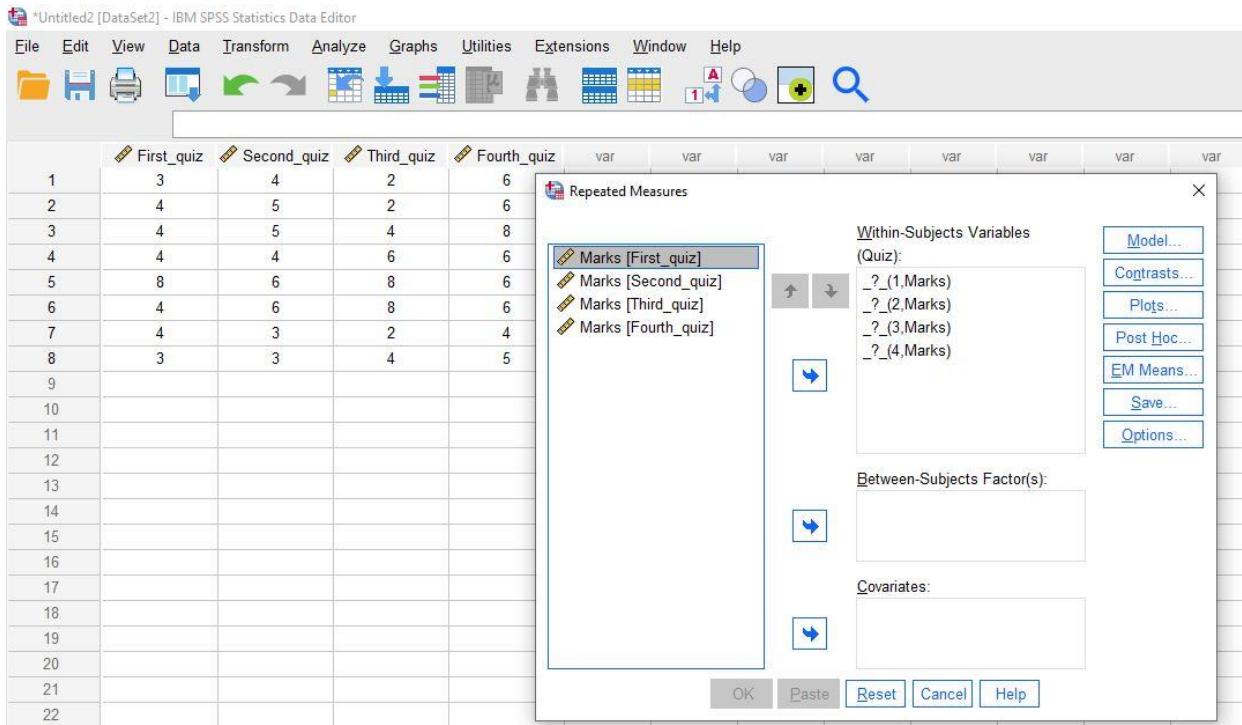
Figure 13.3: Within-subjects variables input (One-way RMA).

The variables corresponding to these levels should be selected and placed in the appropriate space. Select all four variables from the left panel and bring them to the **Within-Subjects Variables** section of the screen (Figure 13.4). There is no **Between-subject factor(s)** in one-way RMA.
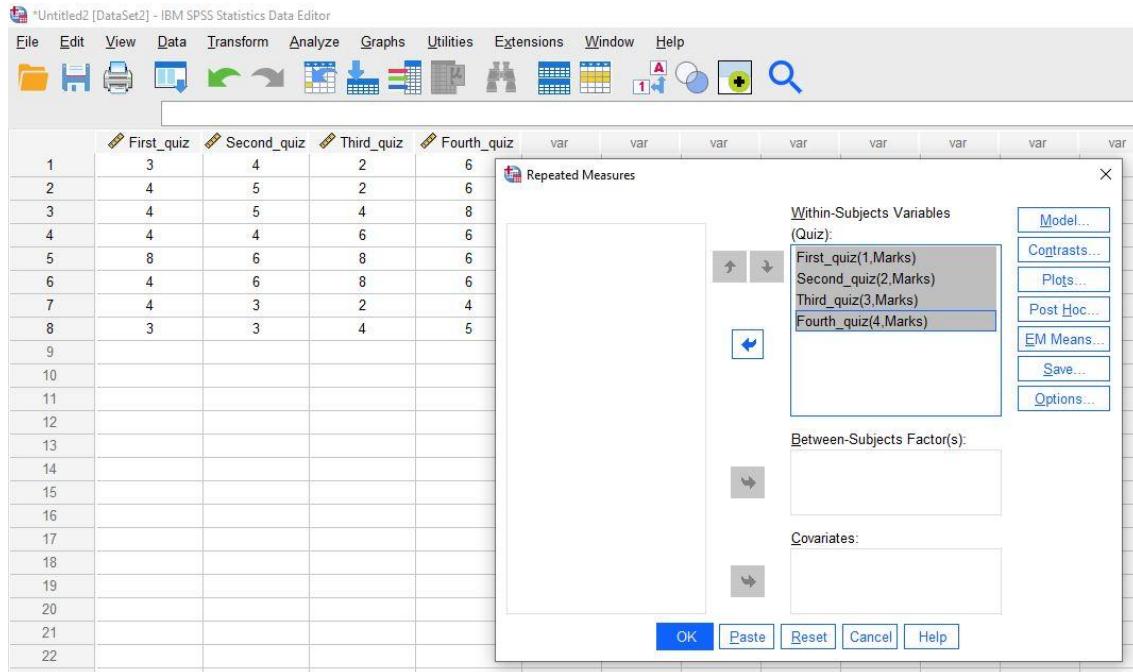
Figure 13.4: Levels selection for within-subject factor (one-way RMA).

All the knowledge belongs to God.

After selecting the variables, first of all, select the **Model** which is suitable for the study plan, in this case **full factorial** is selected as following:
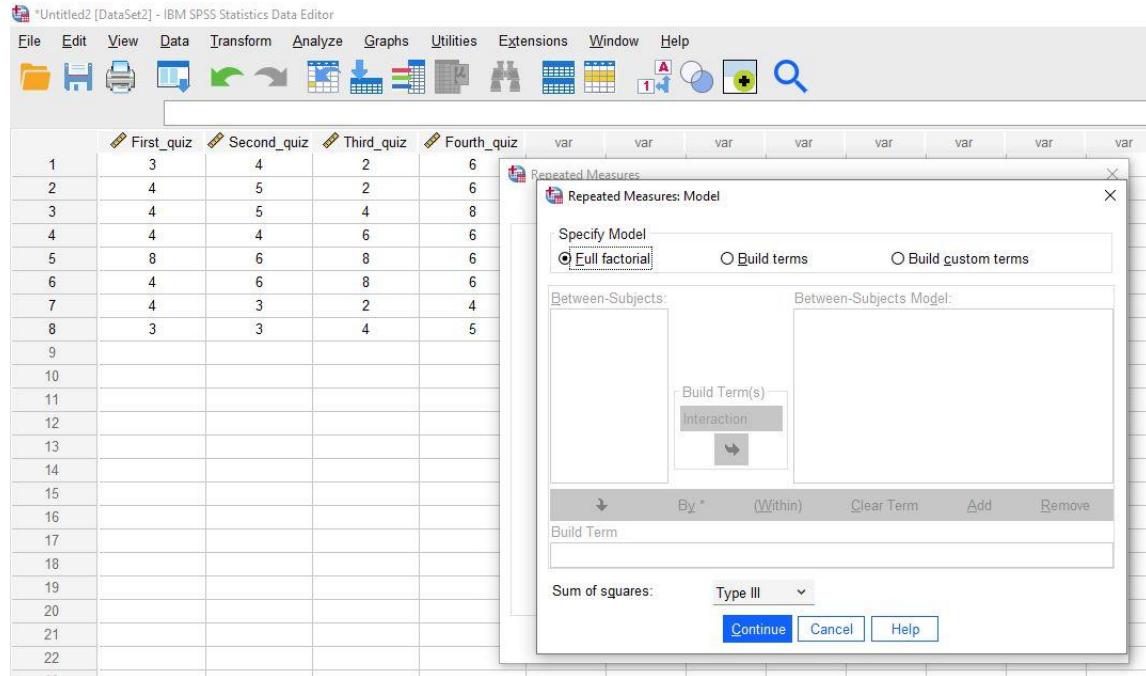
Figure 13.5: Model selection (one-way RMA).

The standard post hoc tests that we have seen for independent designs are not available for repeated-measures analyses (you will find that if you access the post hoc test dialog box it will not list any repeated-measured factors). But researcher can do some basic post hoc procedures through the additional options.

These options can be accessed by clicking on **EM Means** in the main dialog box to open the **Repeated Measures-Estimated Marginal Means** dialog box (Figure 13.6). To specify post hoc tests, select the repeated-measures variable (in this case Quiz) from the box labelled **Estimated Marginal Means**: **Factor(s)** and **Factor** Interactions and drag it to the box labelled **Display Means for**. Once a variable has been transferred, select '**Compare main effect**' (which will now be active). If this option is selected, the box labelled **Confidence interval adjustment** becomes active and you can click on to see a choice of three adjustment levels. The default is to have no adjustment and simply perform a **Tukey LSD** post hoc test (this is not recommended). The second option is a **Bonferroni** correction (recommended for the reasons mentioned above), and the final option is a **Sidak** correction, which should be selected if you are concerned about the loss of power associated with Bonferroni corrected values.

All the knowledge belongs to God.
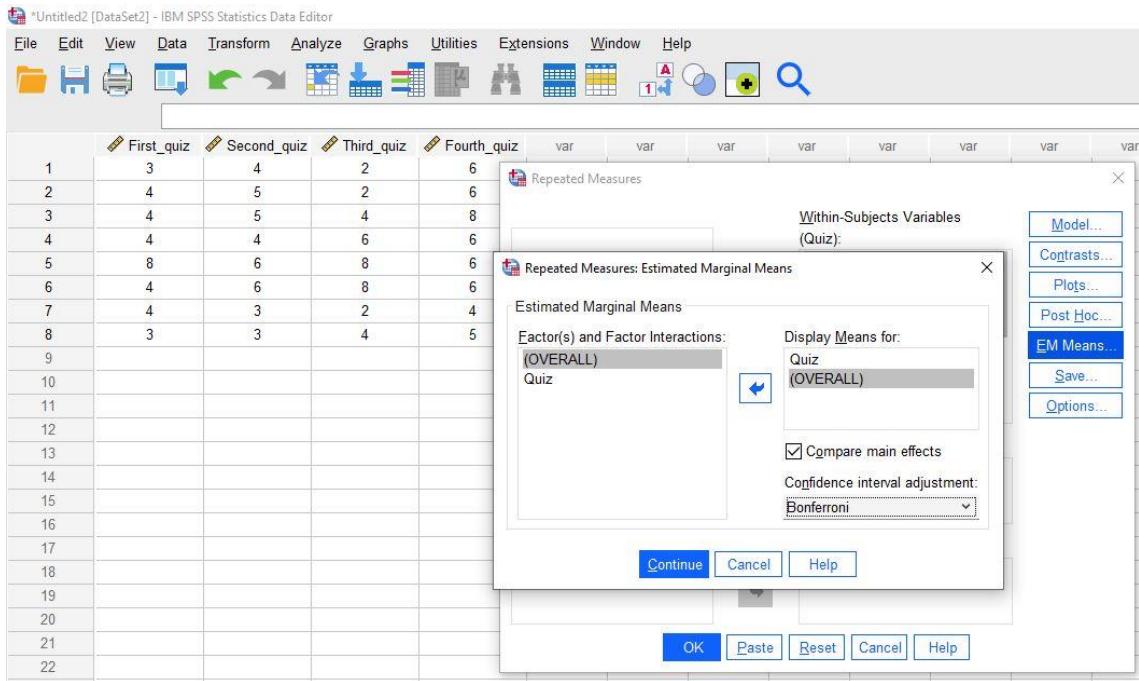
Figure 13.6: Alternative way for Post Hoc (one-way RMA).

Click **Ok** to get the output as following:
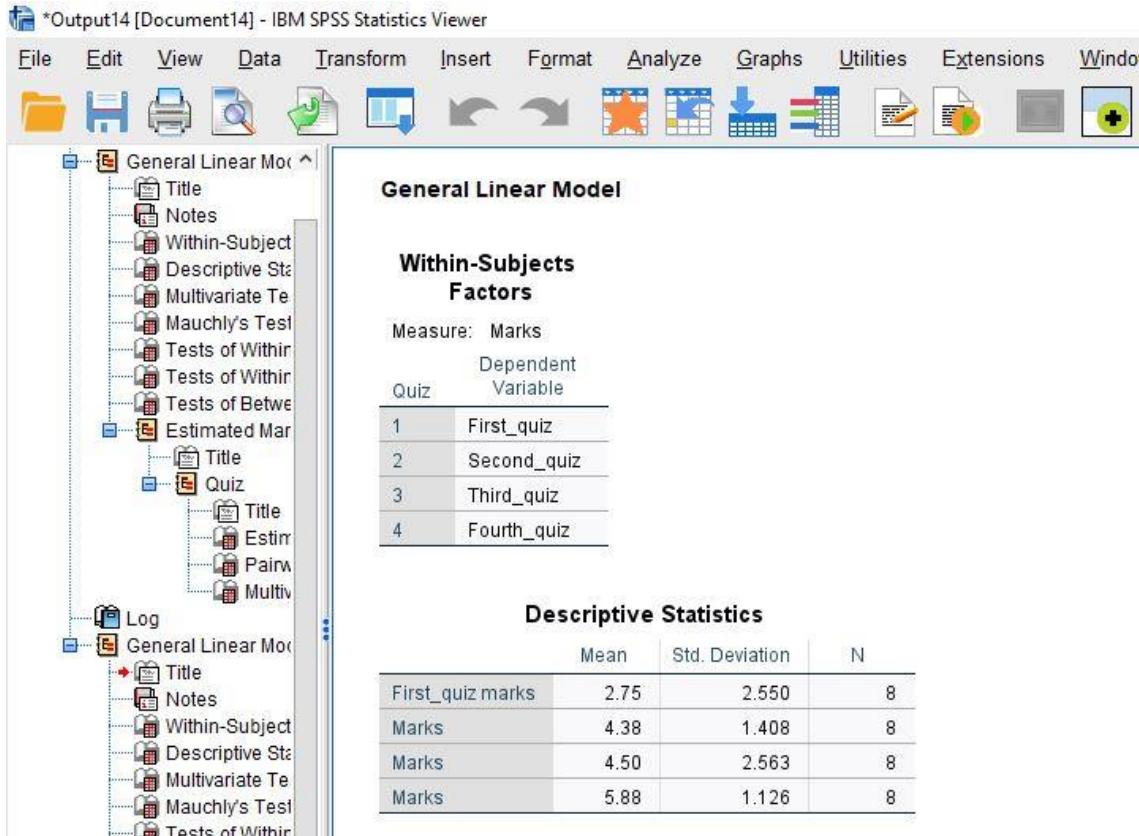
Figure 13.7: Descriptive Statistics output (One-way RMA).

All the knowledge belongs to God.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Quiz	Pillai's Trace	.671	3.406[b]	3.000	5.000	.110
	Wilks' Lambda	.329	3.406[b]	3.000	5.000	.110
	Hotelling's Trace	2.044	3.406[b]	3.000	5.000	.110
	Roy's Largest Root	2.044	3.406[b]	3.000	5.000	.110

a. Design: Intercept
   Within Subjects Design: Quiz

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: Marks

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Quiz	.294	7.006	5	.226	.693	.992	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
   Within Subjects Design: Quiz

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure 13.8: Multivariate Tests and Sphericity output (One-way RMA).

SPSS produces a test known as **Mauchly's tests of Sphericity**, which tests the hypothesis that the variances of the differences between conditions are equal (Figure 13.8). Since Mauchly's test is insignificant, this indicates that the sphericity assumption is sustained. Since sphericity assumption is sustained, no correction is required to be made in the degrees of freedom for the treatment and the error components.

If data violates the sphericity assumption there are several corrections that can be applied to produce a valid F-ratio. SPSS produces three corrections based upon the estimates of sphericity advocated by Greenhouse and Geisser and Huynh and Feldt (Figure 13.8). Both of these estimates give rise to a correction factor that is applied to the degrees of freedom used to assess the observed F-ratio. The calculation of these estimates is beyond the scope of this book; we need know only that the three estimates differ.

A final option, when you have data that violate sphericity, is to use **Multivariate Test Statistics**, because they do not make this assumption (Figure 13.8). However, it is also showing insignificant results, so we don't really need it here.

Tests of Within-Subjects Effects

Measure: Marks

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Quiz	Sphericity Assumed	39.250	3	13.083	3.939	.022
	Greenhouse-Geisser	39.250	2.080	18.871	3.939	.042
	Huynh-Feldt	39.250	2.975	13.191	3.939	.023
	Lower-bound	39.250	1.000	39.250	3.939	.088
Error(Quiz)	Sphericity Assumed	69.750	21	3.321		
	Greenhouse-Geisser	69.750	14.559	4.791		
	Huynh-Feldt	69.750	20.828	3.349		
	Lower-bound	69.750	7.000	9.964		

Tests of Within-Subjects Contrasts

Measure: Marks

Source	Quiz	Type III Sum of Squares	df	Mean Square	F	Sig.
Quiz	Linear	36.100	1	36.100	8.231	.024
	Quadratic	.125	1	.125	.046	.836
	Cubic	3.025	1	3.025	1.050	.340
Error(Quiz)	Linear	30.700	7	4.386		
	Quadratic	18.875	7	2.696		
	Cubic	20.175	7	2.882		

Figure 13.9: Test of Within-Subjects Effects output (One-way RMA).

For the basic repeated-measures ANOVA, we are interested only in the **Tests of Within-Subjects Effects**. It can be seen from the Figure 13.9 that **Sig** value is 0.02 for the **Sphericity Assumed**, as in this case, sphericity assumption sustained (insignificant). Otherwise, two different corrections namely **Greenhouse-Geisser** and the **Huynh-Feldt** are usually applied if the sphericity assumption is violated. Notice that in all cases the F-ratios remain the same; it is the degrees of freedom that change (and hence the critical value against which the obtained F-statistic is compared). The degrees of freedom have been adjusted using the estimates of sphericity calculated in SPSS output.

It can be seen that after applying the **Greenhouse-Geisser** correction, the F value is also significant because associated p value of F is 0.42 which is less than 0.05. In fact, the F is significant in all the situations and no difference in findings occurs due to sphericity assumption. Remember that it is also possible taking an average of the two estimates, and certainly when the two corrections give different results (as is the case here) this can be useful.

The second table displays the contrast output and the linear polynomial is also significant; however, in this case the significant contrast should be ignored.

All the knowledge belongs to God.
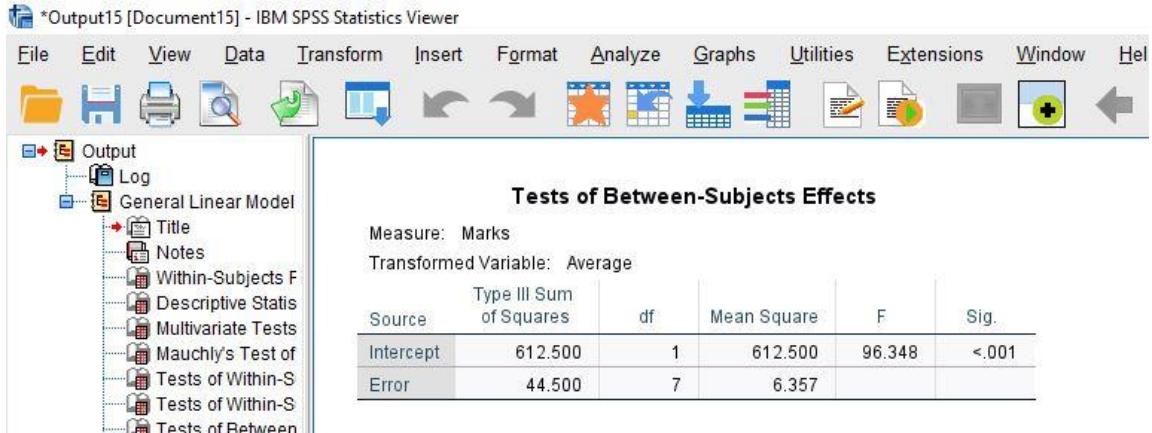
Figure 13.10: Test of Between-Subjects Effects output (One-way RMA).

Although the **Test of Between-Subjects Effects** is significant (Figure 13.10), but we don't need it here. Similarly, in 4th quiz, positive effect has been observed in marks, the pairwise comparison can be taken as basis for description of results among 4 levels:
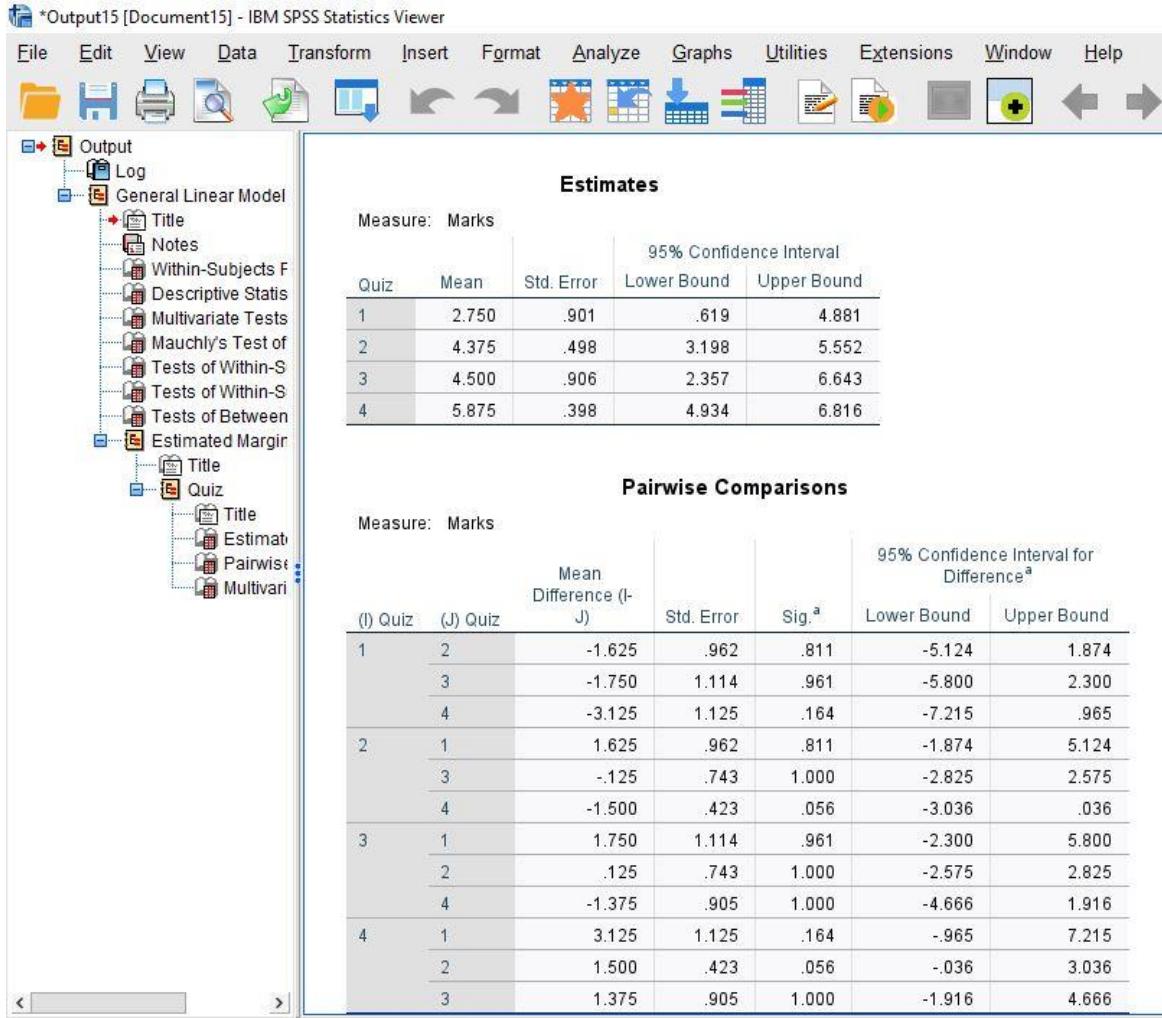
Figure 13.11: Estimated Marginal Means and Pairwise Comparisons output (one-way RMA).

All the knowledge belongs to God.

The last table is **Multivariate Tests**, and it has been explained earlier.

Multivariate Tests					
	Value	F	Hypothesis df	Error df	Sig.
Pillai's trace	.671	3.406[a]	3.000	5.000	.110
Wilks' lambda	.329	3.406[a]	3.000	5.000	.110
Hotelling's trace	2.044	3.406[a]	3.000	5.000	.110
Roy's largest root	2.044	3.406[a]	3.000	5.000	.110

Each F tests the multivariate effect of Quiz. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

Figure 13.12: Multivariate Tests output (one-way RMA).

On the basis of the sampled data, it may be concluded that the teaching efficiency significantly affects the marks of the students with the progression of quiz in time period.

13.3: TWO-WAY RMA

In this type of ANOVA, effect of two factors on some dependent variables is investigated simultaneously where both factors are within-subjects. Since in within-subjects factor all subjects are tested under all treatment conditions, in two-way repeated measures design all subjects are tested in each level of both the factors. If a two-way repeated measures ANOVA is planned where first factor has two levels and the second has three, then in the experiment all the subjects shall be tested under each of the six treatment conditions.

It is also known as **two-factorial ANOVA with repeated measures** or simply **factorial repeated measures ANOVA.** If the design has a between-groups variable and a within-subjects independent variable, it is called a **mixed factorial design.** For example, if the independent variables are gender (a between-groups variable) and time of measurement (with pre-test and post-test as within-subjects levels); this is a 2×2 mixed factorial design with repeated measures on the second factor.

13.3.1: Hypotheses in RMA

Consider an experiment in which a researcher wishes to see the effect of temperature (20°C and 30°C) and exercise machine (treadmill, cycle, and stepper) on sweating loss. Here both factors, temperature and exercise machine, are within-subjects having levels 2 and 3, respectively. In solving this design, the following three types of hypotheses are tested:

- Whether loss of sweat differs in two different temperatures irrespective of the exercise machines
- Whether loss of sweat differs in different exercise machines irrespective of the temperatures
- Whether interaction between temperature and exercise machine is significant

13.3.2: Applications of Two-way RMA

Some of the specific situations where this design can be used by the researchers are as follows:

A sports scientist may like to investigate the effect of warming-up duration (10, 20, and 30 min) and the type of turf (rubberized and cinder) on athlete's performance in an 800-meter event.

All the knowledge belongs to God.

Here warming-up duration and court types are the two independent within-subjects variables having three and two levels, respectively. A group of randomly selected athletes may be tested for their performance on 800-meter event in each of the six treatment conditions. In this design, researchers can test whether the two main effects for warming-up duration and turf type affect the 800-meter performance of athletes. Simultaneously, the researcher can also test the significance of interaction between warming-up and turf type on the 800-meter performance.

Similarly, take another example. In order to develop an appropriate exercise regimen for the people, an exercise scientist may plan a study to investigate the effect of exercise intensity and the temperature on heart rate. In doing so, he may select three exercise intensities (low, medium, and high) and two temperatures (25°C and 30°C) for the study. Thus, a random sample of subjects may be asked to undergo all the six treatment conditions after that they may be tested for their heart rate. This way the effect of exercise, temperature, and their interaction on heart rate may be tested for their significance.

13.3.3: SPSS Procedure (Both Within-Subjects Factors)

A statistician taught a subject, Experimental Design, with different techniques (Modern; Classical; Mixed) to each of 4 Females and 4 Males, PhDs students, to see its impact on grades for assessing teaching efficiency. Marks were obtained on each subject at different quiz which is shown in the Table 13.2. Apply test to report its findings at the significance level 0.05.

Table 13.2: Data on marks (out of 10) obtained on the students.

Gender	Modern	Classical	Mixed
Female	8	2	3
Female	5	2	2
Female	4	0	4
Female	4	0	0
Male	10	2	0
Male	6	0	6
Male	6	2	4
Male	5	3	4

Here it is required to test:

$$H_1 = Change\ in\ teahcing\ technique\ has\ an\ effect\ on\ marks$$

$$H_1' = Change\ in\ gender\ has\ an\ effect\ on\ marks$$

$$H_1' = There\ exist\ an\ interaction\ among\ gender\ and\ technique$$

Variables included in above problem are:
- $Within\ subjects\ factor = Teaching\ technique\ and\ gender$
- $Response = Marks$

Define the variables in **Variable View** as following, because Teaching and Gender, both are within-subject factor (repeated measures), so we will have to make six columns (2×3):

All the knowledge belongs to God.

Figure 13.13: Defining variables (two-way RMA).

In the data file, click the **Data View** and process the following commands in sequence:

Analyze → General Linear Model → Repeated Measures

After clicking **Repeated Measures** option, define the variables (Figure 13.14). By default, the **Within-Subject Factor Name** is written as **factor 1**. Replace the **factor 1** name by the first independent variable Gender and write the number of levels as 2 as there are two levels (Female and Male). Click **Add** to move this information into the box. Similarly, write the name of second independent variable Technique in the **Within-Subject Factor** area and write the number of levels as 3 because it has three levels (Modern, Classical, and Mixed). Click **Add** to move this information into the box. In the **Measure Name** section, type Marks. Click **Add** to move this information into the box.

Figure 13.14: Defining the Within-Subject Factor Name and Measure (two-way RMA).

All the knowledge belongs to God.

Clicking **Define** command in the screen shown in Figure 13.14, it shall take you to the screen as shown in Figure 13.15 for selecting the **Within-Subjects Variables**. Select all 6 variables from the left panel and bring them to the **Within-Subjects Variables** section of the screen:
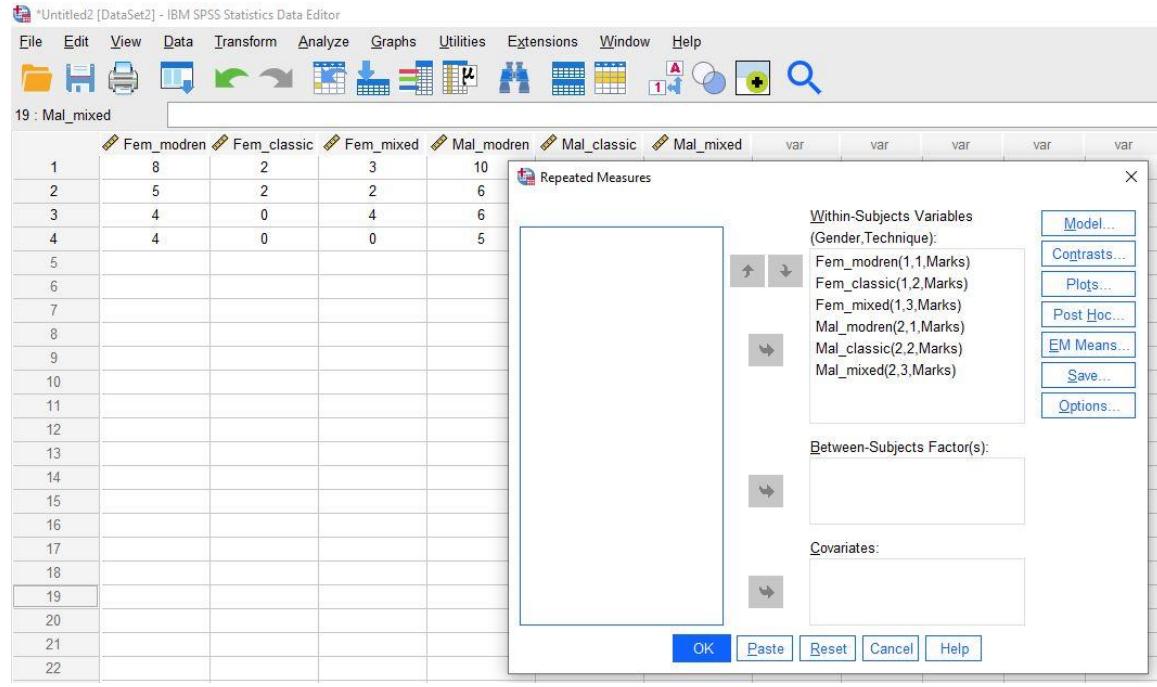
Figure 13.15: Defining Within-Subjects Variables (two-way RMA).

Select the options just like explained in one-way RMA. After selecting variables, click **OK** for generating outputs:
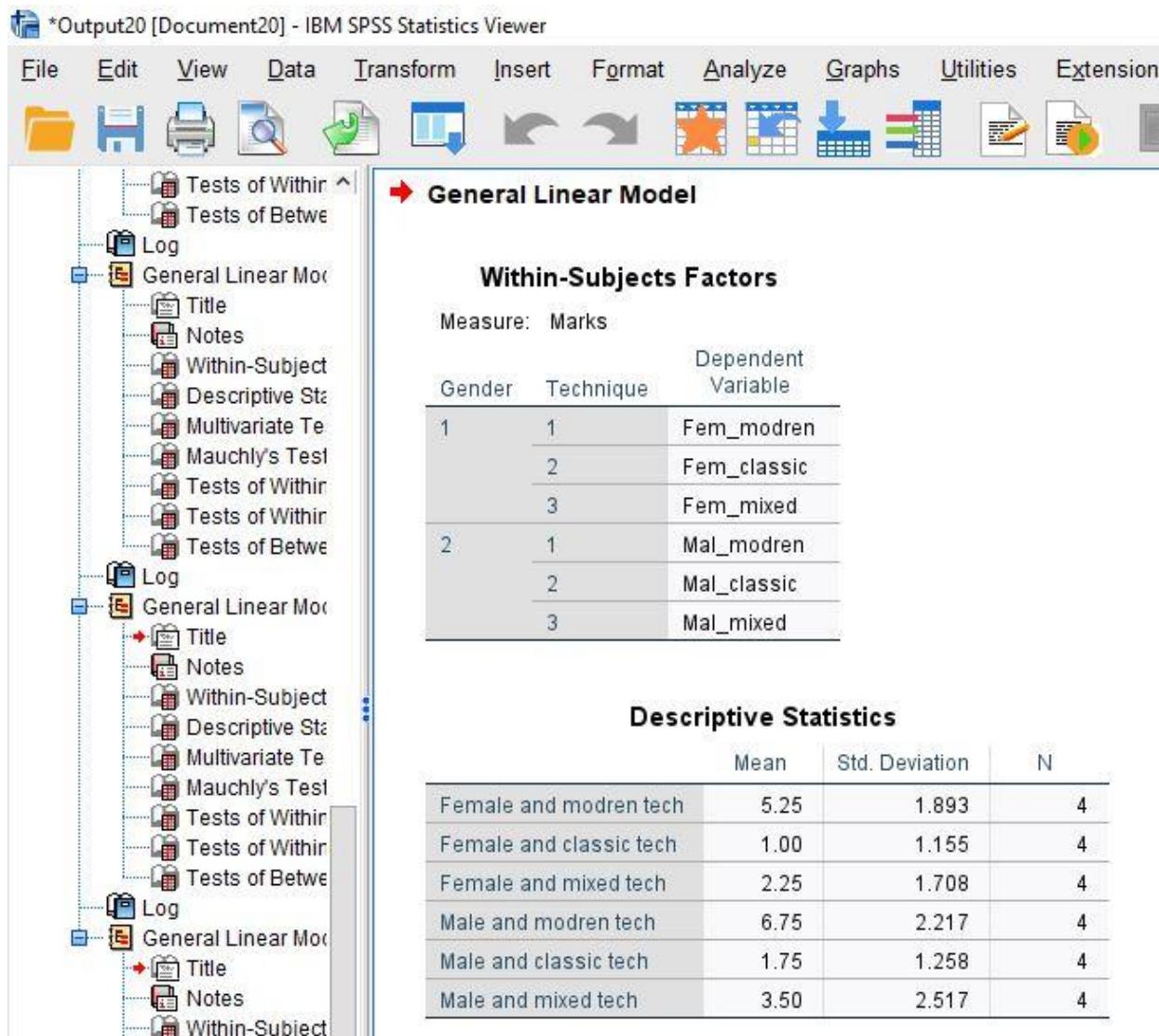
Figure 13.16: Decriptive Statistics output (two-way RMA).

It can be seen from Figure 13.16 that in the Female and Male groups, the marks for the students is high in Modern technique in comparison to that of Classical or Mixed techniques.
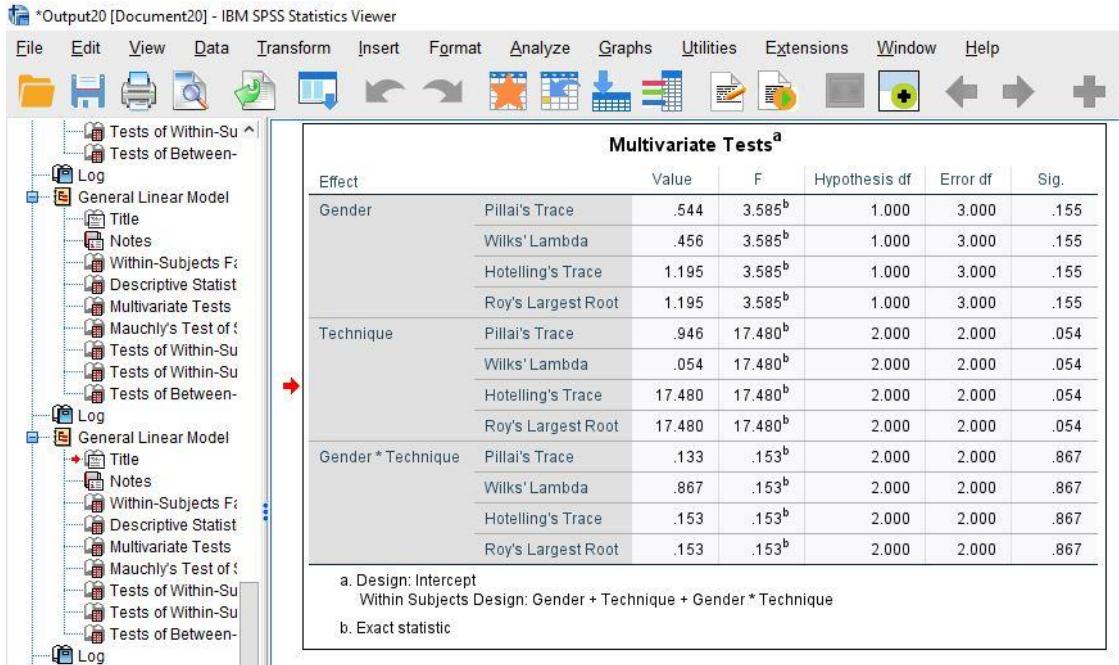
Figure 13.17: Multivariate Tests output (Two-way RMA).

Figure 13.18: Mauchy's test of Sphericity output (Two-way RMA).

It can be seen from the Figure 13.18 that **Mauchly's Test of Sphericity** is insignificant for all because none of the significance value (p value) is less than 0.05. Note that insignificance value of **Mauchly's W** statistic for Gender has not been computed as there is no question of sphericity if the factor has only two levels. Since sphericity assumption is not violated in any of these effects, no correction is required in the degrees of freedom for any of the effect.

All the knowledge belongs to God.

Tests of Within-Subjects Effects

Measure: Marks

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Gender	Sphericity Assumed	8.167	1	8.167	3.585	.155
	Greenhouse-Geisser	8.167	1.000	8.167	3.585	.155
	Huynh-Feldt	8.167	1.000	8.167	3.585	.155
	Lower-bound	8.167	1.000	8.167	3.585	.155
Error(Gender)	Sphericity Assumed	6.833	3	2.278		
	Greenhouse-Geisser	6.833	3.000	2.278		
	Huynh-Feldt	6.833	3.000	2.278		
	Lower-bound	6.833	3.000	2.278		
Technique	Sphericity Assumed	89.083	2	44.542	8.933	.016
	Greenhouse-Geisser	89.083	1.238	71.972	8.933	.042
	Huynh-Feldt	89.083	1.675	53.198	8.933	.024
	Lower-bound	89.083	1.000	89.083	8.933	.058
Error(Technique)	Sphericity Assumed	29.917	6	4.986		
	Greenhouse-Geisser	29.917	3.713	8.057		
	Huynh-Feldt	29.917	5.024	5.955		
	Lower-bound	29.917	3.000	9.972		
Gender * Technique	Sphericity Assumed	.583	2	.292	.095	.911
	Greenhouse-Geisser	.583	1.470	.397	.095	.858
	Huynh-Feldt	.583	2.000	.292	.095	.911
	Lower-bound	.583	1.000	.583	.095	.778
Error(Gender*Technique)	Sphericity Assumed	18.417	6	3.069		
	Greenhouse-Geisser	18.417	4.409	4.177		
	Huynh-Feldt	18.417	6.000	3.069		
	Lower-bound	18.417	3.000	6.139		

Figure 13.19: Tests of Within-Subjects Effects output (Two-way RMA).

Although Teaching Technique is significant (main effect); however, interaction effect is insignificant ($P > 0.05$) (Figure 13.19) so there exist no interaction between gender and teaching technique. On the basis of the sampled data, it may be concluded that the marks of students are not affected in comparison to that of gender due to the teaching technique.
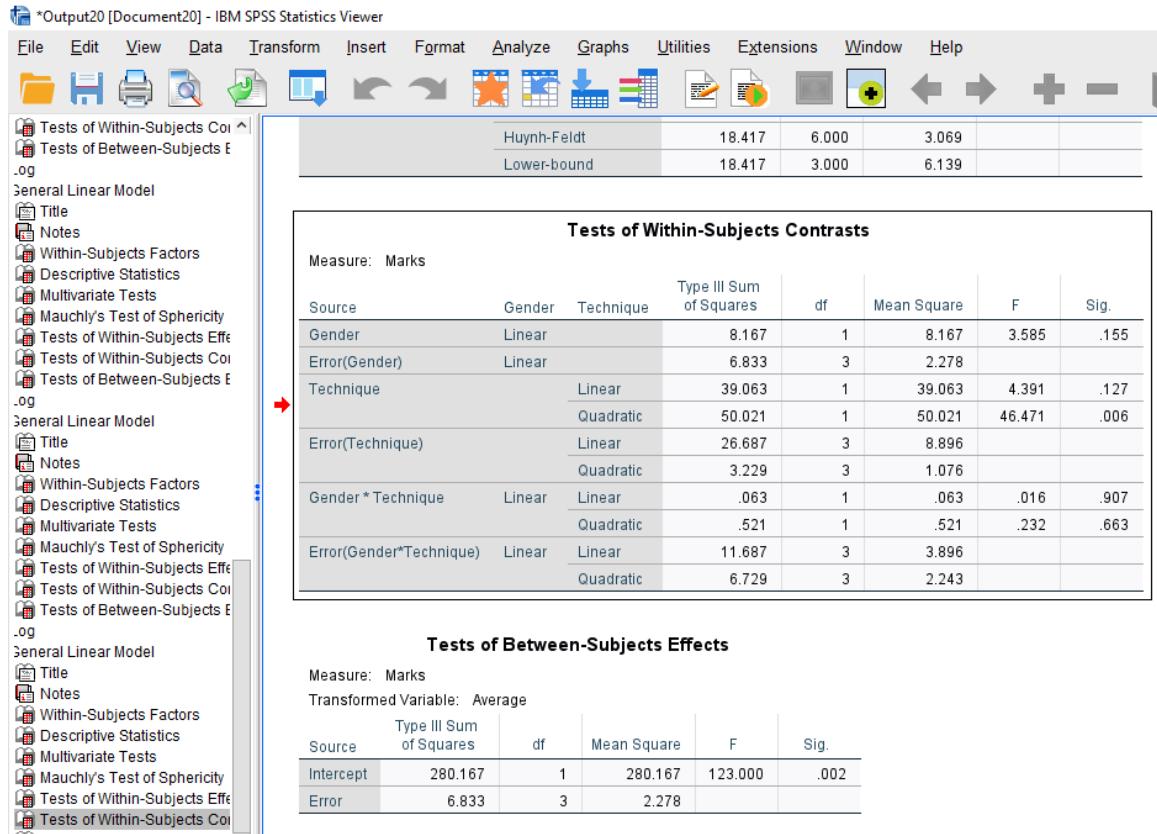
Figure 13.20: Tests of Between-Subjects Effects output (Two-way RMA).

Note that in case of repeated measures design, no post hoc test is applicable because the data is obtained by repeated measures. It is because of this reason that SPSS does not show the option for post hoc when variables are within-subjects.

13.3.4: SPSS Procedure (1 Within-subjects and 1 Between-subjects Factor)

Suppose we want to check the effect of a diet and gender on chicken weight gain (Table 13.3) at time zero (week 0); at time one (after 2 weeks); at time two (after 4 week) with 95% confidence level.

Table 13.3: Weight gain (g) of female and male chicken.

	Time		
	Week 0	**Week 2**	**Week 4**
Female	25	30	57
	24	40	50
	22	40	65
Male	20	30	35
	20	28	40
	24	28	39

There will be at least three hypotheses as:

$$H_1 = Change\ in\ time\ has\ an\ effect\ on\ weight\ gain$$

All the knowledge belongs to God.

$$H'_1 = Change\ in\ gender\ has\ an\ effect\ on\ weight\ gain$$

$$H''_1 = There\ exist\ an\ interaction\ among\ gender\ and\ time\ on\ response$$

Define the variables in **Variable View** as following.

Figure 13.21: Defining variable (two-way RMA).

The **Values** for Gender variable must be coded as 1 and 2 for female and male respectively (Figure 13.21). Insert data into SPSS module and use commands:

$Analyze \rightarrow GLM \rightarrow Repeated\ meaures\ ANOVA$

Figure 13.22: Two-way RMA data input (1 between and 1 within-subject factor).

All the knowledge belongs to God.

You can understand from the SPSS module that within-subjects are always in columns and between-subjects are always in rows in SPSS table (Figure 13.22). **Within-subject Factor Name** will be Time and **Number of Levels** will be 3. Similarly, other factor (Gender) will be **Between-Subject Factor(s)**. The Weight_gain will be **Measure Name**.

After inserting the information, the user would have to explain levels in **Define** tab (Figure 13.22) which will be open a new screen as in Figure 13.23.
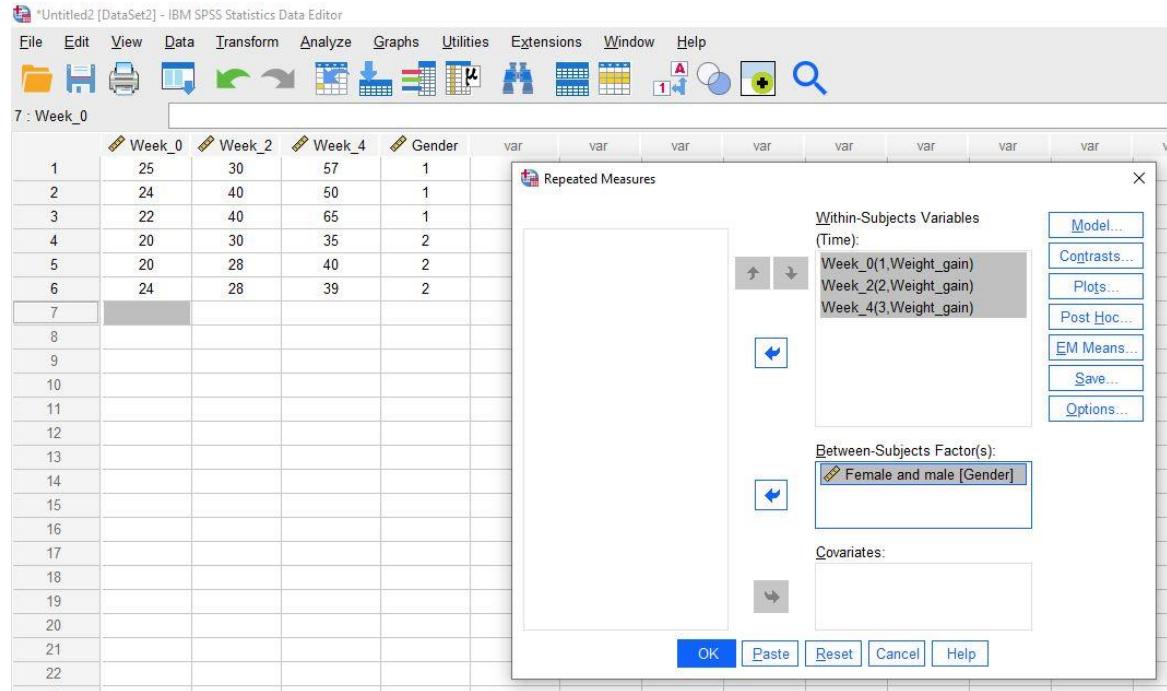
Figure 13.23: Defining the Within- and Between-Subjects Factors (Two-way RMA).

If there is any covariant in study (say age), it can be inserted into **Covariates** tab (Figure 13.23). Otherwise, select gender for **Post-hoc** and other options from **Option** tab and click **OK** to get output:
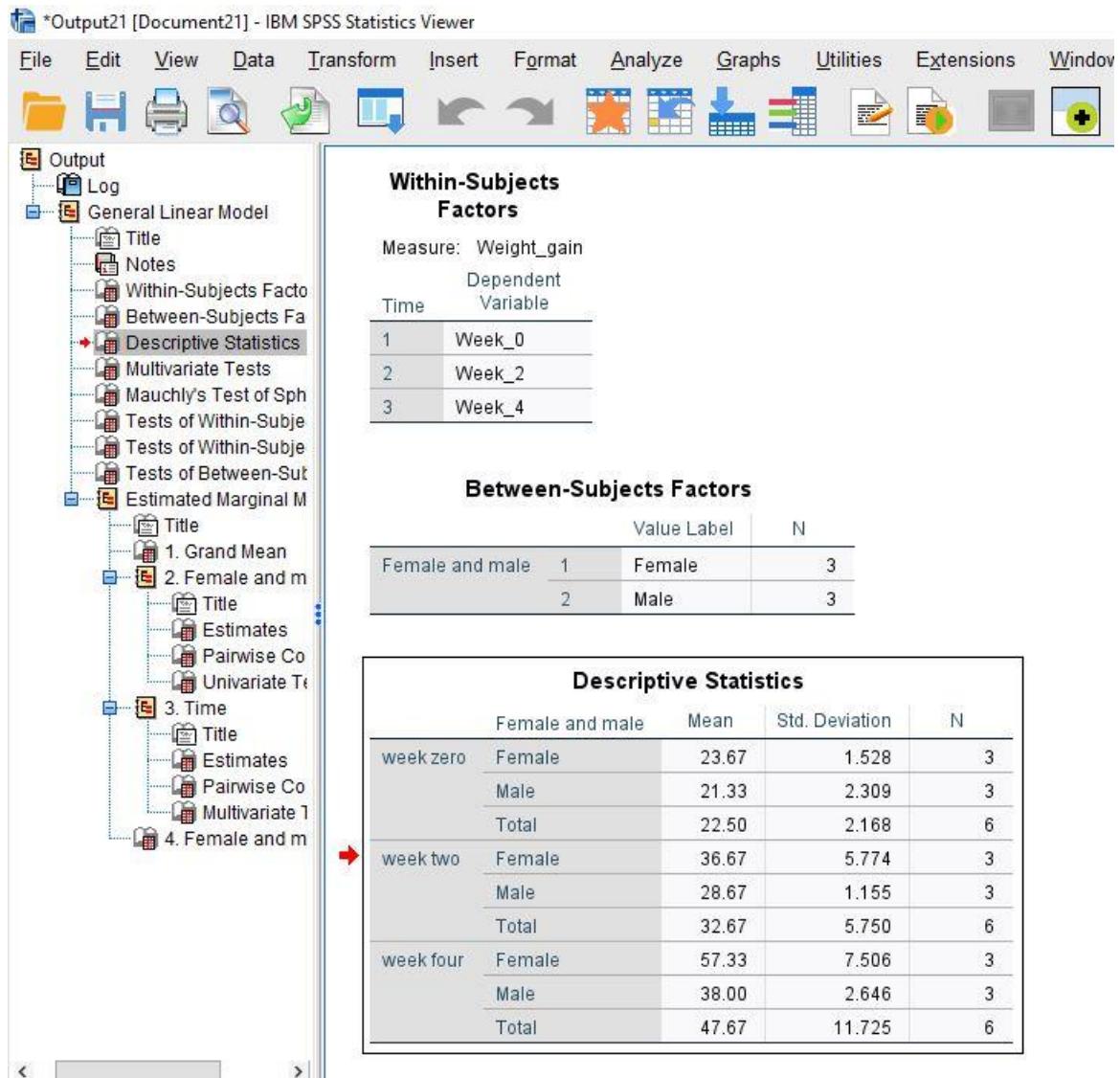
Figure 13.24: Descriptive Statistics output (Two-way RMA).

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Time	Pillai's Trace	.960	35.844[b]	2.000	3.000	.008
	Wilks' Lambda	.040	35.844[b]	2.000	3.000	.008
	Hotelling's Trace	23.896	35.844[b]	2.000	3.000	.008
	Roy's Largest Root	23.896	35.844[b]	2.000	3.000	.008
Time * Gender	Pillai's Trace	.726	3.982[b]	2.000	3.000	.143
	Wilks' Lambda	.274	3.982[b]	2.000	3.000	.143
	Hotelling's Trace	2.654	3.982[b]	2.000	3.000	.143
	Roy's Largest Root	2.654	3.982[b]	2.000	3.000	.143

a. Design: Intercept + Gender
   Within Subjects Design: Time

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: Weight_gain

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	.907	.292	2	.864	.915	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + Gender
   Within Subjects Design: Time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure 13.25: Multivariate and Munchy's test output (Two-way RMA).

Munchy's Test of Sphericity is actually the extension of Levene's test table. The significance of this table (Figure 13.25) indicates whether the sphericity (plural of variances) is assumed or not. In this case **Sig** is 0.864 which means equal variances are assumed (H_1 is rejected).

Tests of Within-Subjects Effects

Measure: Weight_gain

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Time	Sphericity Assumed	1923.444	2	961.722	47.754	<.001
	Greenhouse-Geisser	1923.444	1.830	1050.915	47.754	<.001
	Huynh-Feldt	1923.444	2.000	961.722	47.754	<.001
	Lower-bound	1923.444	1.000	1923.444	47.754	.002
Time * Gender	Sphericity Assumed	224.778	2	112.389	5.581	.030
	Greenhouse-Geisser	224.778	1.830	122.812	5.581	.035
	Huynh-Feldt	224.778	2.000	112.389	5.581	.030
	Lower-bound	224.778	1.000	224.778	5.581	.077
Error(Time)	Sphericity Assumed	161.111	8	20.139		
	Greenhouse-Geisser	161.111	7.321	22.007		
	Huynh-Feldt	161.111	8.000	20.139		
	Lower-bound	161.111	4.000	40.278		

Tests of Within-Subjects Contrasts

Measure: Weight_gain

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	Linear	1900.083	1	1900.083	91.940	<.001
	Quadratic	23.361	1	23.361	1.191	.336
Time * Gender	Linear	216.750	1	216.750	10.488	.032
	Quadratic	8.028	1	8.028	.409	.557
Error(Time)	Linear	82.667	4	20.667		
	Quadratic	78.444	4	19.611		

Figure 13.26: Test of Within-Subject Factors output (Two-way RMA).

The only table needs to be elaborated in results is **Tests of Within-Subjects Effects** (Figure 13.26). Tests are mentioned in **Source** column, anyone test can be taken (which favors our result) for explaining out result. But remember to use only the first one (**Sphericity Assumed**) one test for all the effects, as sphericity is assumed in this case.

Both main effect (Time) and Interaction effect are significant which means H_1 is accepted.
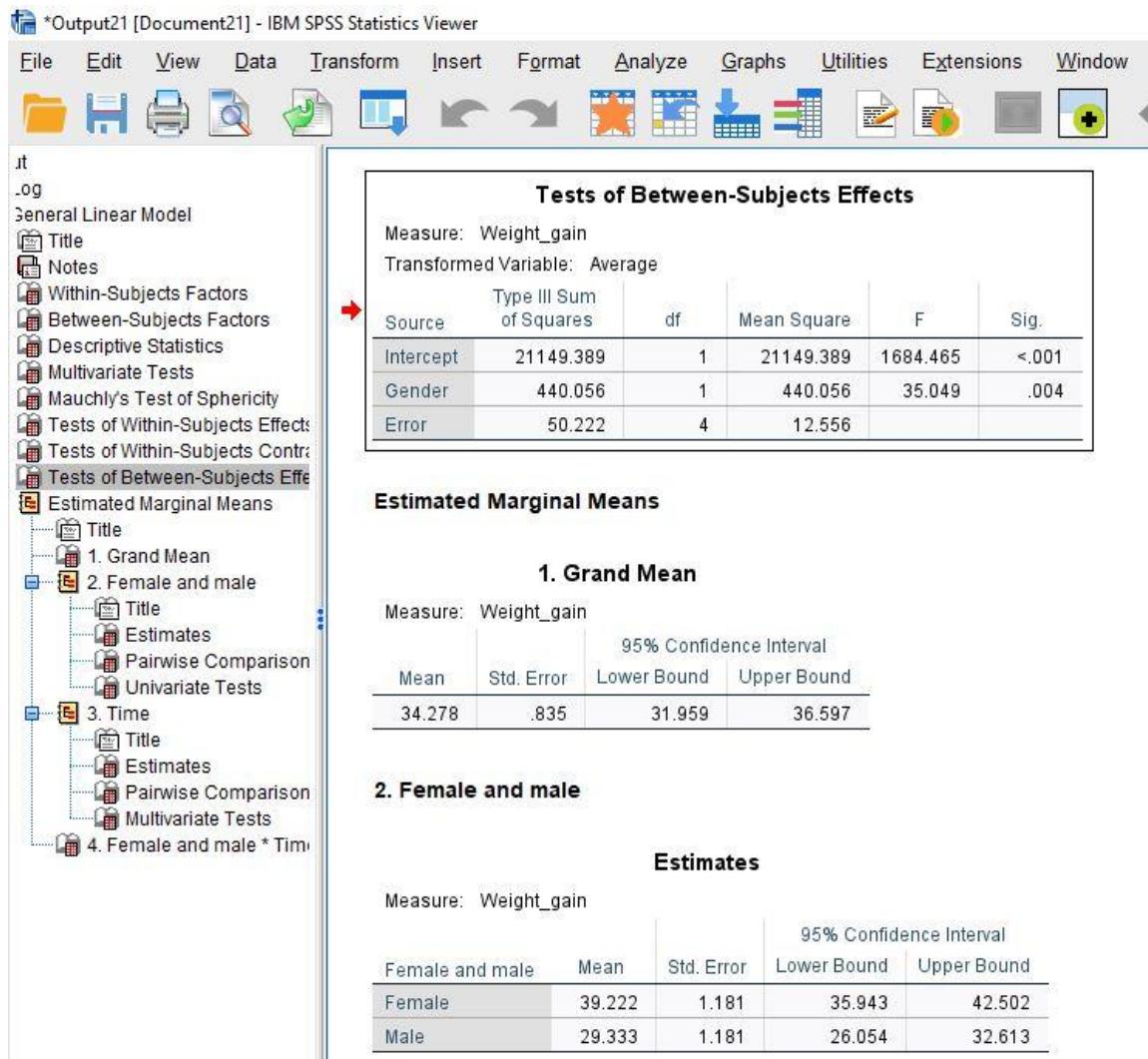
Figure 13.27: Test of Between-Subjects Effects and Estimates output (Two-way RMA).

Tests of Between-Subjects Effects for factor (Gender) is also significant (Figure 13.27). However, we will focus on interaction only.

2. Female and male

Estimates

Measure: Weight_gain

Female and male	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Female	39.222	1.181	35.943	42.502
Male	29.333	1.181	26.054	32.613

Pairwise Comparisons

Measure: Weight_gain

(I) Female and male	(J) Female and male	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Female	Male	9.889*	1.670	.004	5.251	14.527
Male	Female	-9.889*	1.670	.004	-14.527	-5.251

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Univariate Tests

Measure: Weight_gain

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	146.685	1	146.685	35.049	.004
Error	16.741	4	4.185		

The F tests the effect of Female and male. This test is based on the linearly

Figure 13.28: Pairwise Comparisons, female-male (two-way RMA).

Pairwise Comparisons (Gender) also show that there is positive improvement when we move from female to male gender in time (Figure 13.28). Similarly, **Pairwise Comparison** (Time) can be further clarified from following tables (Figure 13.29). There is a positive development in progression of time.

All the knowledge belongs to God.

3. Time

Estimates

Measure: Weight_gain

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	22.500	.799	20.281	24.719
2	32.667	1.700	27.948	37.386
3	47.667	2.297	41.288	54.045

Pairwise Comparisons

Measure: Weight_gain

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-10.167*	2.205	.030	-18.899	-1.434
	3	-25.167*	2.625	.002	-35.562	-14.771
2	1	10.167*	2.205	.030	1.434	18.899
	3	-15.000*	2.896	.020	-26.472	-3.528
3	1	25.167*	2.625	.002	14.771	35.562
	2	15.000*	2.896	.020	3.528	26.472

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Figure 13.29: Pairwise Comparison, time (two-way RMA).

The interaction among time and gender can also be judged from the following table as well:

4. Female and male * Time

Measure: Weight_gain

Female and male	Time	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Female	1	23.667	1.130	20.528	26.805
	2	36.667	2.404	29.993	43.340
	3	57.333	3.249	48.313	66.354
Male	1	21.333	1.130	18.195	24.472
	2	28.667	2.404	21.993	35.340
	3	38.000	3.249	28.980	47.020

Figure 13.30: Gender and Time interaction (two-way RMA).

All the knowledge belongs to God.

Post Hoc (Duncan) is not possible as only 2 levels are for gender (between-subject factor). It can be concluded that the weight gain is significantly improved with respect to time and the female has more pronounced effect of weight gain as compared to the males.

CHAPTER 14: ANALYSIS OF COVARIANCE

14.1: COVARIATE

Sometimes there are certain variables which may or may not affect the research outcome and are not part of research, but they are directly related to the study variable. Such are Co-variables. These are predictive responses, meaning that covariates are responses measured for an experimental unit in anticipation that the covariates will be associated with, and thus predictors for, the primary response.

Analysis of covariance (ANCOVA) design is used when randomization of treatments is not possible, and the treatments are allocated to the intact group. It allows to remove the effect of a known covariate. Sometimes, a researcher may not have a choice to assign treatments randomly and may be compelled to administer treatments to intact groups that are not homogeneous. The subjects in these intact groups may differ initially on many parameters, and therefore statistical control is necessary to reduce the experimental error due to such initial differences in groups. Thus, in experimental research, the individual variations that appear within the measures on the dependent variable are potentially correlated with something else.

Suppose if the dependent variable is a measure of how well the subjects learn swimming under one or the other of the two methods of instructions, the potential correlates are likely to include such parameters as prior relevant learning, endurance, strength, motivation, self-discipline, intelligence, etc. These potential correlates are known as covariates, typically interval-level variables. That is why, applying one-way ANOVA on the data obtained by taking the difference of post- and pre-testing in all the treatment group results in wrong conclusion. This is because treatment effect is not compensated due to initial variations among the groups.

ANCOVA can be used to compare the effectiveness of these instructional methods on learning swimming after removing the effect of the identified covariates. Similarly, if in an experiment with a response variable 'y' there is another variable, say 'x', and that y is linearly related to x. And that x cannot be controlled by the experimenter but can be observed along with y. The variable x is called a covariate or **concomitant variable.**

ANCOVA is always used for more than 2 independent means when there is a covariant. It typically is used to adjust or control for differences between the groups based on another covariate. Imagine that we found that boys and girls differ on math achievement. However, this could be due to the fact that boys take more math courses in high school. ANCOVA allows us to adjust the math achievement scores based on the relationship between number of math courses taken and math achievement. We can then determine if boys and girls still have different math achievement scores after making the adjustment.

ANCOVA also requires that the functional relationship of the covariates with the character of primary interest is known beforehand. Consider the case of a rice variety trial in which weed incidence is used as a covariate. With a known functional relationship between weed incidence and grain yield, the character of primary interest, the covariance analysis can adjust grain yield in each plot to a common level of weed incidence. With this adjustment, the variation in yield due to weed incidence is quantified and effectively separated from that due to varietal difference.

Analysis of covariance can be applied to more complex treatment structures, such as factorial designs (**factorial experiments with covariates**). Provided enough data exists for every treatment combination, nearly any complex treatment structure can be analyzed through the analysis of covariance approach.

14.1.1: Covariance versus Blocking

Although blocking and covariance technique are both used to reduce experimental error, the differences between the two techniques are such that they are usually not interchangeable. The

All the knowledge belongs to God.

ANCOVA, for example, can be used only when the covariate representing the heterogeneity between experimental units can be measured quantitatively. However, that is not a necessary condition for blocking. In addition, because blocking is done before the start of the experiment, it can be used only to cope with sources of variation that are known or predictable. ANCOVA, on the other hand, can take care of unexpected sources of variation that occur during the experiment. Thus, covariance analysis is useful as a supplementary procedure to take care of sources of variation that cannot be accounted for by blocking.

Blocking is highly effective when the pattern of variation between experimental units is recognizable in terms of a set of distinct groups. If potential variation between units corresponds to qualitative differences between units, then blocking is natural. Covariance allows adjustment for as many covariate factors as are thought appropriate. If the form of the covariance relationship is properly identified, then a covariance adjustment for variation between units caused by quantitative differences between units will produce more precise results than the approximation which blocking offers to a continuously varying pattern of yield over the available units.

14.2: APPLICATIONS OF ANCOVA

Consider a hypothetical experiment having four varieties where each variety is replicated three times. Some of the varieties in some of the plots are assumed to be affected by rodents. Particularly variety number 2, which performed very well in the first two replications, performed poorly in the last replication. This shows some varieties performed below their true genetic potential due to factors not initially considered. What will be the result if we proceed with the calculation of ANOVA using the observed data set? It is obvious that the comparison will definitely be in favor of the varieties not affected by rodents. Thus, the result shows non-significance among the four varieties despite a clear mean difference in the two replications. However, we have to construct another variable which can be used as an adjustment factor. For this particular data, we assumed that a three-scale rating of the yield condition is sufficient. Scale number 1 is given to the severely-damaged plot. Scale number 2 for plots with moderate damage, and 3 for plots not damaged at all. When this variable is used as a covariate, the result of ANOVA is completely changed since the difference among the treatments is significant at about 3% probability level.

In another scenario, US wanted to make such type of bombs which could penetrate in earth to certain depth to target underground constructions of Iran's nuclear facilities. But she does not know how much their depth is. So, they took help from satellite images to know the amount of soil debris they dig out to guess the amount of depth. In this scenario, the penetration power of bomb is the main factor (study variable) they want to test, but the amount of soil debris is a covariance.

Similarly, suppose cows of different ages have an effect on milking frequency and so these cows cannot be taken as same one entity. Ages are covariant here. That is why covariant is taken as **random variable** (uncontrolled). Note that the covariance is different from the local control which is related to lab conditions only like the availability of different types of instruments or incubators etc.

An experiment is conducted to study the effect of two different types of exercises, that is, A and B on the cardiovascular efficiency of the subjects. Further, an experimenter is forced to use three intact groups of subjects from three different colleges. However, there

All the knowledge belongs to God.

is a freedom to assign treatments (types of exercises) randomly to the groups. Out of the three groups, one may serve as control. Since the treatments cannot be randomly assigned to the subjects, the possibility of initial differences (before administration of treatment) among the groups on their cardiovascular efficiency exists. Thus, one may decide to have initial measurement (X) of cardiovascular efficiency on each subject before applying treatment. This measure of covariate, to be measured before implementing the exercise, is used to adjust measurements on the cardiovascular efficiency (Y) obtained after the treatment. Thus, the variables X and Y can be defined as follows:

- X is pre-treatment scores of the cardiovascular efficiency in each of the three treatment groups
- Y is post-treatment scores of the cardiovascular efficiency in each of the three treatment groups

ANCOVA can be further elaborated from another situation where it is decided to start a conditioning program for the students to improve their physical efficiency. Three conditioning programs (C1, C2, and C3) having different intensities and durations have been suggested by the experts. A researcher needs to decide as to which program should be implemented. For this, he conducts an experiment on university students in which he allocates C1 treatment on the undergraduate, C2 on the graduate, and C3 on the research students. Since these groups differ in their age, variation in their post-treatment performance on the physical efficiency may be partly due to variation in their age. Here age (X) is a covariate and post-treatment data on post-treatment performance (Y) is a dependent variable. The ANCOVA design may be used to compare the effectiveness of these conditioning programs by compensating the variation in age.

Can diet affect the size of a caterpillar's head? Such an effect is plausible, because a caterpillar's chewing muscles occupy a large part of the head. To study the effect of diet, a biologist raised caterpillars (Pseudaletia unipuncta) on three different diets: diet 1, an artificial soft diet; diet 2, soft grasses; and diet 3, hard grasses. He measured the weight of the head and of the entire body in the final stage of larval development. The results are shown in Figure 14.1, where Y= ln (head weight) is plotted against X= ln (body weight), with different symbols for the three diets. Note that the effect of diet is striking; there is virtually no overlap between the three groups of points.
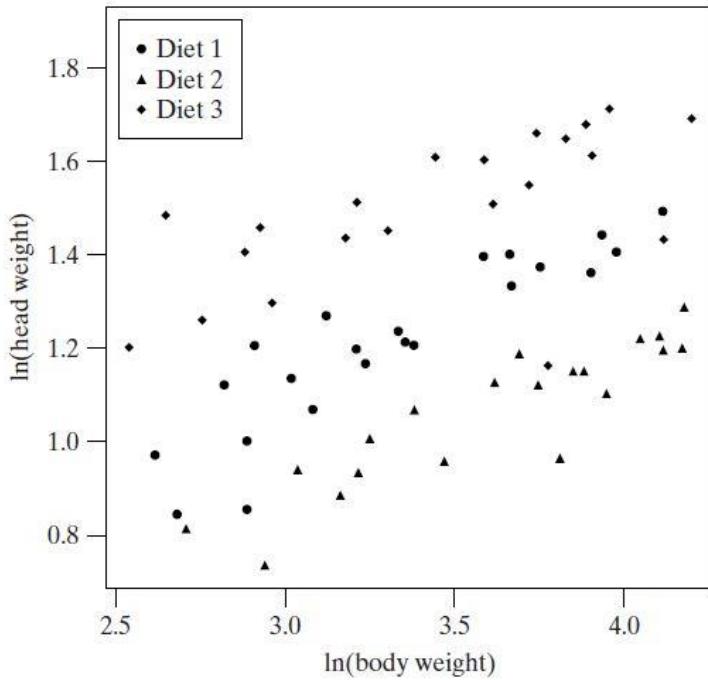
Figure 14.1: Weight of the head and entire body in the final stage of larval development.

But if we were to ignore X and consider Y only, as displayed in Figure 14.2, the effect of diet would be much less pronounced. This shows how comparison of several groups with respect to a variable Y can be strengthened by using information on an auxiliary variable X that is correlated with Y. A classical method of statistical analysis for such data is analysis of covariance.
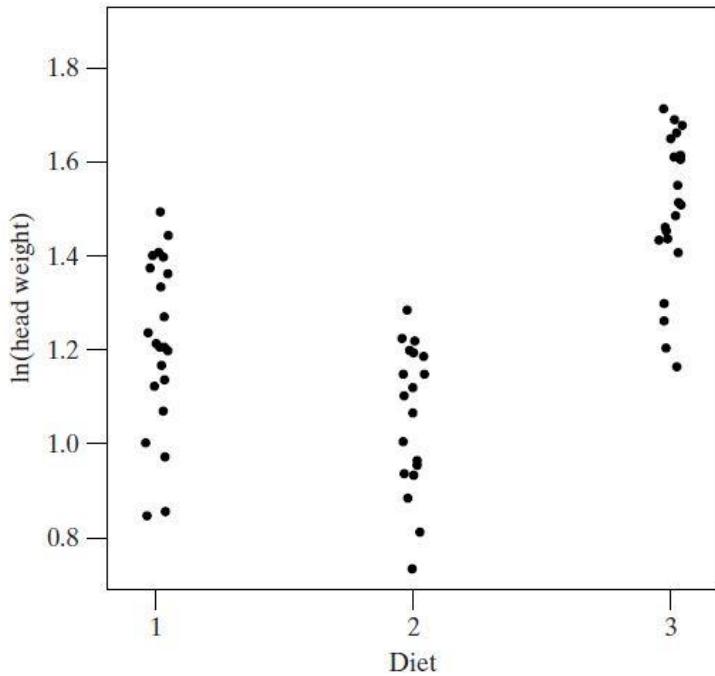
Figure 14.2: Weight of the head under diet influence only.

All the knowledge belongs to God.

14.3: PROS AND CONS OF COVARIATE

We can make comparisons between treatments more precise by including covariates in our model and this is not completely separate from the first advantage—covariate models allow us to compare predicted treatment responses at a common value of the covariate for all treatments. So, the use of ANCOVA is to reduce within-group error variance by allowing the covariate to explain some of this error variance. In any experiment, there may be unmeasured variables that confound the results (i.e., variables that vary systematically with the experimental manipulation). If any variables are known to influence the dependent variable being measured, then ANCOVA is ideally suited to remove the bias of these variables. Once a possible confounding variable has been identified, it can be measured and entered into the analysis as a covariate (elimination of confounding).

However, one potential pitfall of covariate models is that they assume that the covariate is not affected by the treatment. Remember that co-variants are always used for need or necessity.

14.4: ASSUMPTION FOR ANCOVA

For ANCOVA there is a fourth assumption, linearity, that there is a linear relationship between the covariates and the dependent variable. This can be checked with a scatterplot (or matrix scatterplot if there is more than one covariate). The regression slopes for the covariates (in relation to the dependent variable) need to be the same for each group (this is called homogeneity of regression slopes). This assumption is one of the most important assumptions, and it can be checked with an F-test on the interaction of the independent variables with the covariate. If the F-test is significant, then this assumption has been violated.

14.5: MODEL OF ANCOVA

Covariate models have two parts: a usual treatment effect part and a covariate effect part. The treatment effect part is essentially determined by the design, as usual; but there are several possibilities for the covariate effect part, and our model will be appropriate for the data only when we have accurately modeled the relationship between the covariates and the response. Remember that ANCOVA works only when >80% variables are quantitative among all the research variables. But there should be no more than 1-2 covariant in study.

$$Response = Study\ variables\ (qualitative) + Quantitative$$

14.6: SPSS PROCEDURE

Suppose we have different diets to check the effect on weight gain of chickens (Table 14.1). But there is an assumption that ages (3 weeks and 4 weeks) may have effect on the outcome and here comes the concept of covariance.

All the knowledge belongs to God.

Table 14.1: Weight gain of chicken under diet and age influence.

	Diet A	Diet B
Day 21	40	65
	44	55
	55	60
	45	98
Day 28	43	88
	38	88
	70	88
	65	90

Variables included in this problem are:

- $Dependent\ variable = weight\ gain$
- $Fixed\ factor = diet$
- $Covariate = age$

Our hypothesis will be:

$H_1 = At\ least\ one\ of\ the\ diets\ have\ different\ response\ form\ others$

Here two variables, namely Diet and Age, need to be defined along with their properties. Diet is a scale variable, whereas Treatment is a nominal variable. These variables along with their characteristics can be defined in the **Variable View** as following:
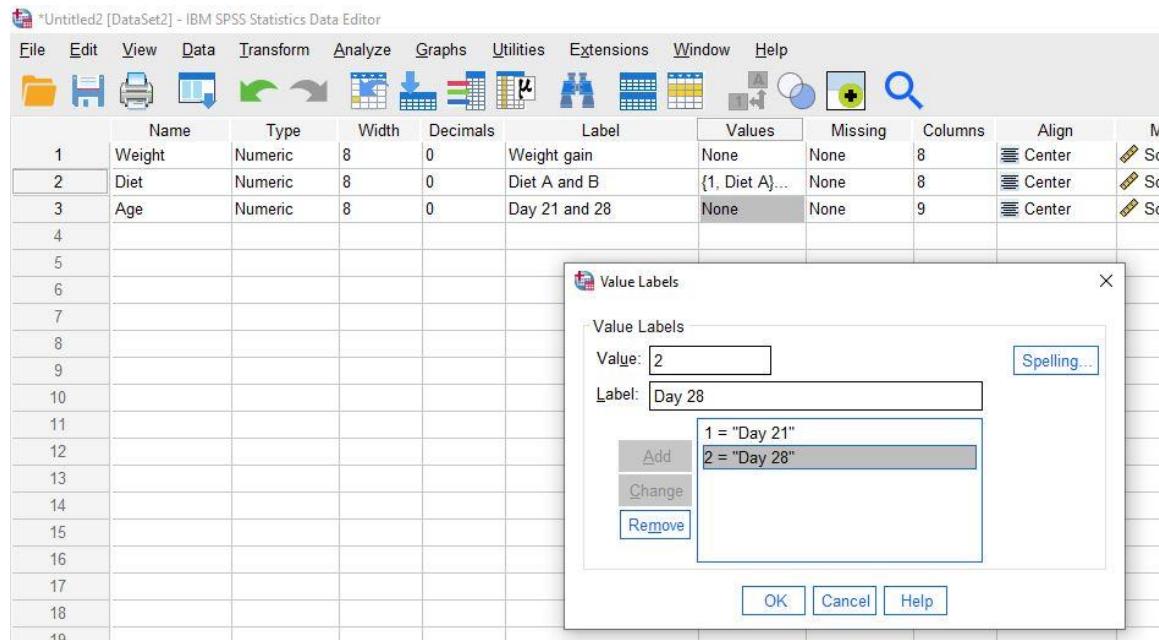
Figure 14.3: Defining variable in ANCOVA (SPSS).

Remember to assign coding in **Values** for Diet and Age variables (Figure 14.3). Now we will have weight as dependent variable, and ages as covariance. After inserting data into SPSS, in **Data View**, click on the following commands in sequence:

$Analyze \rightarrow General\ Linear\ Model \rightarrow Univariate$

All the knowledge belongs to God.
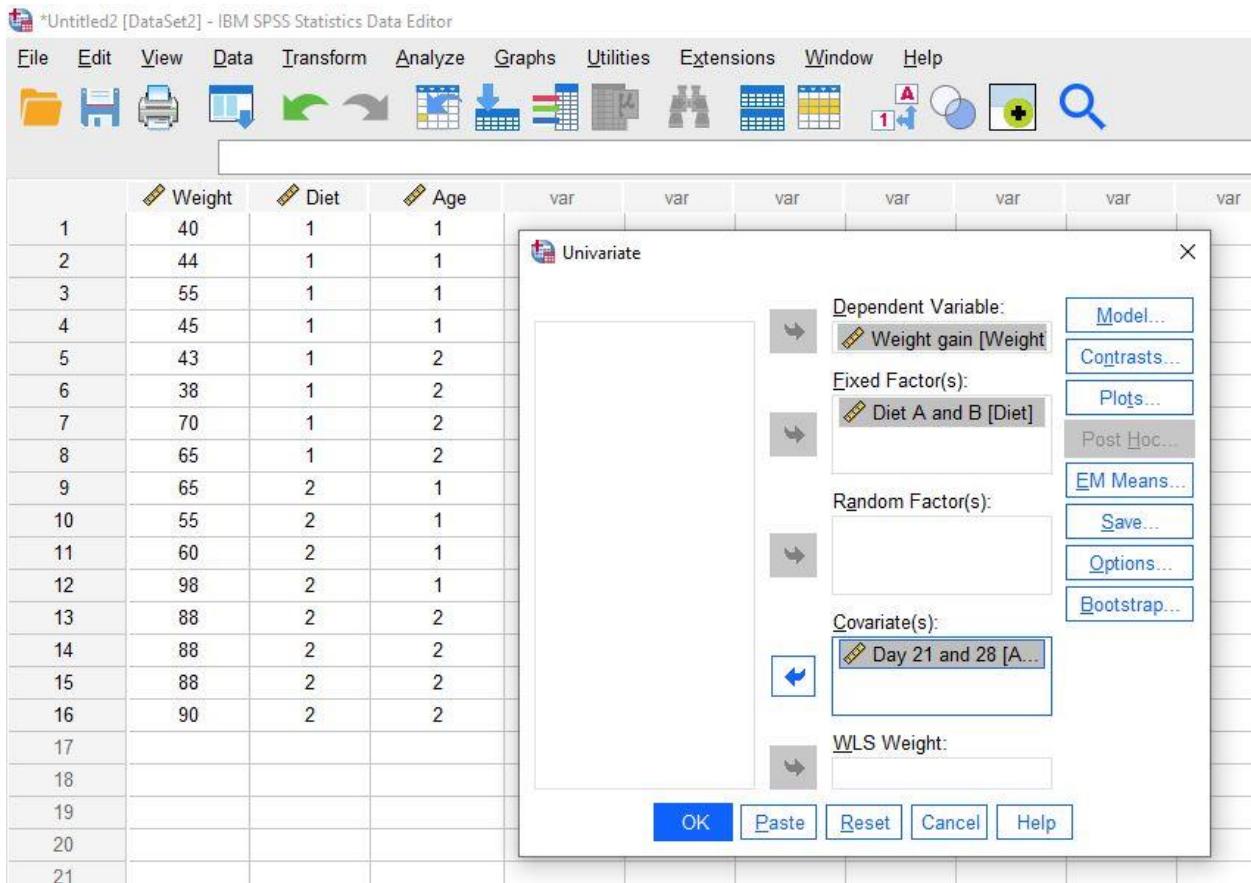
Figure 14.4; Variable and covariate input (ANCOVA).

We have diet as **Fixed Factor(s)** and weight gain as **Dependent Variable**. Age will be **Covariate(s)** (Figure 14.4). Note that **Post Hoc** option is disabled and cannot be applied here. After selecting the confidence level from **Options**, the output will be:
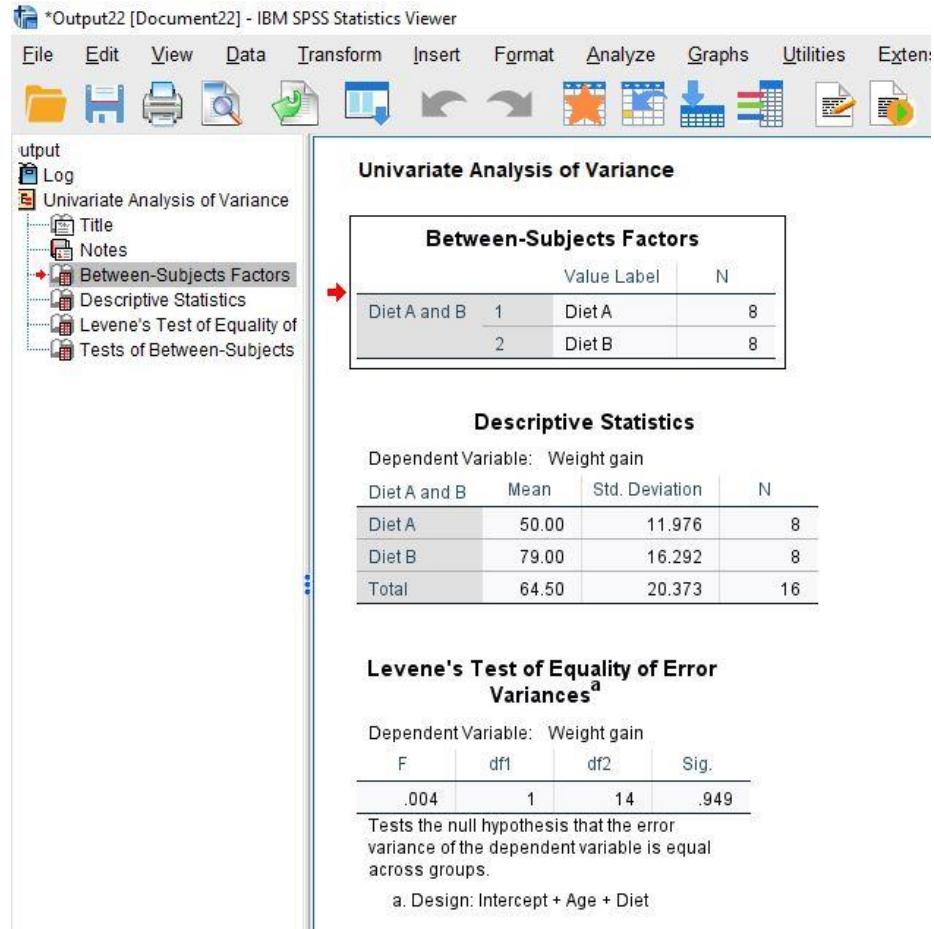
Figure 14.5: Descriptive Statistics and Leven's Test output (ANCOVA).

The **Levene's test** is insignificant (Figure 14.5), so the assumption of equal variance has not been violated.

Tests of Between-Subjects Effects

Dependent Variable: Weight gain

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4093.000[a]	2	2046.500	12.473	<.001
Intercept	3132.900	1	3132.900	19.094	<.001
Age	729.000	1	729.000	4.443	.055
Diet	3364.000	1	3364.000	20.503	<.001
Error	2133.000	13	164.077		
Total	72790.000	16			
Corrected Total	6226.000	15			

a. R Squared = .657 (Adjusted R Squared = .605)

Figure 14.6: Tests of Between-Subjects Effects output (ANCOVA).

All the knowledge belongs to God.

Note that there is no interaction between Age and Diet as Age has been taken as covariant (Figure 14.6). The Age is insignificant here ($P = 0.05$), so there is no effect of Age on weight gain of chickens and it cannot be ignored. If we do not use Age as covariance, there will be no change in result of diet effect. However, our focus is Age here which is insignificant.

CHAPTER 15: CROSSOVER DESIGNS

15.1: INTRODUCTION

One particular form of any basic design (CRD, RCBD or LSD etc.) has **time as one blocking factor** with the different treatments applied in sequence to each experimental unit, the treatment sequences being different for different units. The experiment will include a number of patients, subjects or animals who each receive different treatments in each of a number of successive periods. Essentially the experimental unit is redefined to be an observation for an individual subject in a short period of time.

If an experiment involves biological entities, such as animals or humans, since such entities typically exhibit rather larger variability. One way to reduce natural variability is to use each animal or human, which we shall refer to as subjects, as a block rather than as an experimental unit (Table 15.1). Different treatments are then applied successively, that is, in different time periods, to each subject, so that each **subject–period combination** now represents the experimental unit. This is often referred to as **each subject acting as its own control**, since now comparisons between treatments can be made within rather than between subjects.

Following is a multiple Latin square, or cross-over design for two treatments:

Table 15.1: Time and subjects as blocking factor in crossover design.

		Subject									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Period	**1**	A	B	B	A	A	B	A	B	B	A
	2	B	A	A	B	B	A	B	A	A	A

These designs are applied when EUs are normally cheap and easily available, because they require certain number of runs/cycle for an experiment as per treatments. They are used in a situation where individuals (subjects) are used as one blocking factor and time period as the other blocking factor. These designs have been used extensively in different kinds of experimental settings, but mainly in the pharmaceutical industry during the testing of new drugs, in animal science for feeding experiments, and in psychological studies. Suppose the doctor changes the potency of a medicine for the same patient after some time. It will be a cross-over design. These are most efficient but cost is prohibitive factor that is why these are rarely used practically.

The basic idea is that each individual receives (sequentially) all or some of the treatments, one at any given time period, and that for different individuals the order of the application of the treatments is being changed. In other words, individual subjects receive multiple treatments over the course of experiment, with each subject receiving each treatment an equal number of times. As these trials generally have fewer subjects. These designs employ blocking in two directions: spatial blocking is accomplished by grouping subjects into groups to promote within-group factor and second is time period blocking. Essentially it is argued that time is treated as a factor, resulting in a second level of experimental units.

Even though the designs for different applications have the same features, they are often referred to by different names, such as **change-over designs, carry-over designs, switch-over designs, counter-balanced designs**, and sometimes more generally and generically by **repeated measurement designs**. Because each subject serves its own control, crossover design is useful for increasing power and saving the number of subjects need via elimination of response variation between subjects when treatments are compared.

All the knowledge belongs to God.

In such a design participants would be randomly assigned to one of the two treatments sequences (**two-period crossover design**) (Figure 15.1):
- Control period followed by the intervention period.
- Intervention period followed by the control period.
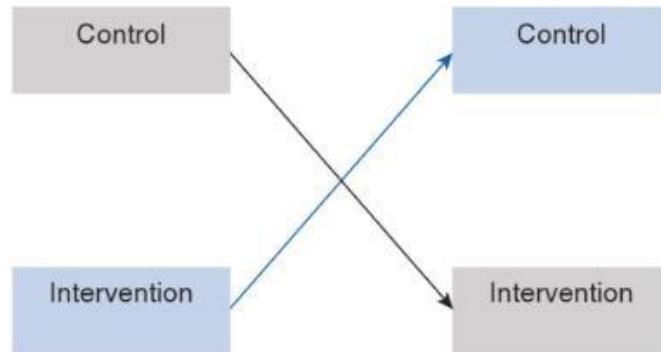
Figure 15.1: Illustration of conditions in two-period crossover design.

One group will receive the intervention for a set period of time and then the intervention will be withdrawn. The other group of subjects will have a period without treatment, and they will then crossover to the intervention condition. All of the participants will receive the treatment at some time. Similarly, there can be **three-period crossover designs** and more.

In the simplest case, where there are two kinds of treatments A and B, the units are first grouped into two pairs like in randomized block design. Suppose from previous knowledge one unit in each pair is expected to give a higher response than the other, and the difference is in favor of the superior unit, it is expected to be about the same in all pairs. It is advisable to ensure that half the replicate of each treatment is applied to the 'poorer' members, and the other half to the 'better' members. The pairs of replicates, which must be even in number, are divided at random into two equal sets—the first set to receive treatment A and the second treatment B.

15.1.1: Crossover versus RMA

Although within-subjects factor and the repetition of treatments may be present in both of these designs; however, the major difference among these two is sequence of treatments in given period of time. In RMA, all the treatments are given to subjects at one time and response is taken and the subjects may receive treatments more than one time. But in crossover designs, some treatments are given to few subjects for given time and then other treatments are given for another time period; however, the treatments are not repeated. In this way the carryover effect is removed from the subjects.

15.1.2: CRD as Cross-over Design

Suppose if a design have 3 levels of single factor and is run once. It is now a simple CRD. Because it must be run two more times with different sequence of levels for design to be cross over. Meanwhile the EUs in first run should not be in the same treatment as were in early runs (Table 15.2). That is why it is not a new design but the extension of early designs i.e., a CRD design with one factor (say diet) and 3 levels (1, 2 and 3) for the weight

gain effect on chicken would need to be run 3 times with such a configuration that the EUs in first run under 1ˢᵗ level should be under 2ⁿᵈ or 3ʳᵈ during next run.

Table 15.2: CRD as cross-over design.

Factor	Diet 1	Diet 2	Diet 3
1ˢᵗ run	A	B	C
2ⁿᵈ run	B	C	A
3ʳᵈ run	C	A	B

A, B and C are experimental units and note that for each run the different sequence is ensured to avoid biasness. However, for this one cross-over design, three CRD runs have been completed. Similarly, for 3 factors each with 3 levels, we need 9 runs for cross-over design. It is the same as old designs, but we need large amount of sample size which are easily available. It is different from LSD as no time period segregation is possible in LSD but is possible in cross-over design. Remember that cross-over designs are usually used in CRD and RCBD but not for LSD and GLSD.

15.1.3: LSD as Cross-over Design

One of the more common uses for a Latin Square arises when a sequence of treatments is given to a subject over several time periods. We need to block on subjects, because each subject tends to respond differently, and we need to block on time-period, because there may be consistently different crossover design over time due to growth, aging, disease progression, or other factors. A crossover design has each treatment given once to each subject and has each treatment occurring an equal number of times in each time period.

For example, a human performance analyst is studying the effect of two replacement fluids on dehydration in 16 subjects (Table 15.3). In the first period, half of the subjects (chosen at random) are given fluid A and the other half fluid B. At the end of the period, the response is measured and a period of time is allowed to pass in which any physiological effect of the fluids is eliminated. Then the experimenter has the subjects who took fluid A take fluid B and those who took fluid B take fluid A. It is a crossover. It is analyzed as a set of 8 Latin squares with two rows (time periods) and two treatments (fluid types). The two columns in each of the 8 squares correspond to subjects.

Table 15.3: LSD as crossover.

Subject																
	I		II		III		IV		V		VI		VII		VIII	
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Period 1	A	B	B	A	B	A	B	B	A	B	B	A	A	B	A	B
Period 2	B	A	A	B	A	B	A	A	B	A	A	B	B	A	B	A

Notice that the rows in the Latin square represent the time periods and the columns represent the subjects. The 8 subjects who received fluid A first (1, 4, 6, 7, 9, 12, 13, 15, 17, and 19) are randomly determined.

All the knowledge belongs to God.

15.1.4: Pros and Cons of Crossover Design

In many experiments, it is found that the variance of observations for different patients or animals may be more than ten times greater than the variation between observations at different times for the same patient or animal. This suggests that enormous gains in precision can be achieved by observing more than one treatment on each patient or animal. A further advantage is that fewer subjects are required. This improvement in statistical precision might seem to provide an overwhelming argument for the use of crossover. However, there are other statistical considerations than precision in the design of experiments, and, in particular, the questions of validity and of the population for which the experimental results are relevant must be considered.

The difficulty with the crossover designs is that the conclusions are appropriate to units similar to those included in the experiment, i.e., to subjects to which treatments are given for a short time period in the context of a sequence of different treatments. We then have to ask if the observed difference between two treatments would be expected to be the same if the same treatment is applied consistently to each subject to which it is allocated. This is a problem of the interpretation of the results from an experiment to subsequent use, and it is a problem which must be considered carefully for all experiments. It is particularly acute in the case of cross-over designs because the experiment is so different from subsequent use. Furthermore, a major disadvantage may arise if the treatments exhibit effects beyond the period in which they are applied. These lingering effects are referred to as residual effects or carryover effects.

15.1.5: Conditions for Crossover Design

Obviously, for this procedure to be of any value certain conditions have to be fulfilled:

- A subject reacts to the treatment soon after the treatment has been applied
- The treatment effect only lasts for a limited time, after this time the subject is restored to its original state
- The treatment effects are the same in each period.

If these conditions are satisfied, we may use some form of block design where the subjects are the blocks and the treatments to be administered to the subject are applied at random. If, however, period effects are suspected, that is, the subjects change systematically over the time of the trial, then we employ some sort of row–column design, with rows representing the periods and columns representing the subjects. This situation may occur, for example, in a dairy cattle feeding trial, where the treatments are applied during the cow's lactation period and where it is known that changes occur during the lactation period regardless of the treatments.

15.1.6: Model of Crossover Design

The model of crossover is the same as basic designs except that the 'time' is mentioned in the model:

$$Response = \ Factor\ (s) + time + Error$$

15.2: SPSS PROCEDURE

Experiment was conducted to determine how the Teaching Method (Modern, Classic or Mixed) affected the marks of students (Table 15.4). Twelve PhD students were recruited for the study, and in order to compare the teaching method effect within each subject, all students learnt all three teaching methods. Following design was used.

All the knowledge belongs to God.

Table 15.4: Teaching method design for students.

	Period I			Period II		
Group	**Modern**	**Classic**	**Mixed**	**Modern**	**Classic**	**Mixed**
A	2	4	2	8	4	0
	3	0	2	4	4	2
B	6	2	0	8	2	2
	6	2	0	6	2	2
C	5	1	2	6	0	2
	5	2	2	6	0	2

Two students were randomized to each sequence group in each treatment. Those students who learnt by modern teaching method, learnt by classic next time and so on. Each student group only learnt by a method one time in experiment. The data resulting from this experiment is the marks (1-10) for each student and is shown in following table. Analyze the findings.

15.2.1: Solution

Variables included in this problem are:

- $Factor\ 1 = teaching\ method$
- $Constraint\ 1 = time\ as\ blocking\ (time\ period)$
- $Constraint\ 2 = students\ as\ blocking\ (groups)$

It is a crossover LSD, and the hypotheses will be:

$$H_1 = Change\ in\ teaching\ method\ has\ an\ effect\ on\ marks$$

$$H_1' = Change\ in\ time\ period\ has\ an\ effect\ on\ marks$$

$$H_1'' = Change\ in\ group\ has\ an\ effect\ on\ marks$$

$$H_1'' = There\ exist\ an\ interaction\ among\ teaching\ method, time\ period\ and\ groups$$

SPSS procedure is the same as three-way ANOVA. First of all, define the variables in the **Variable View** as following:

All the knowledge belongs to God.
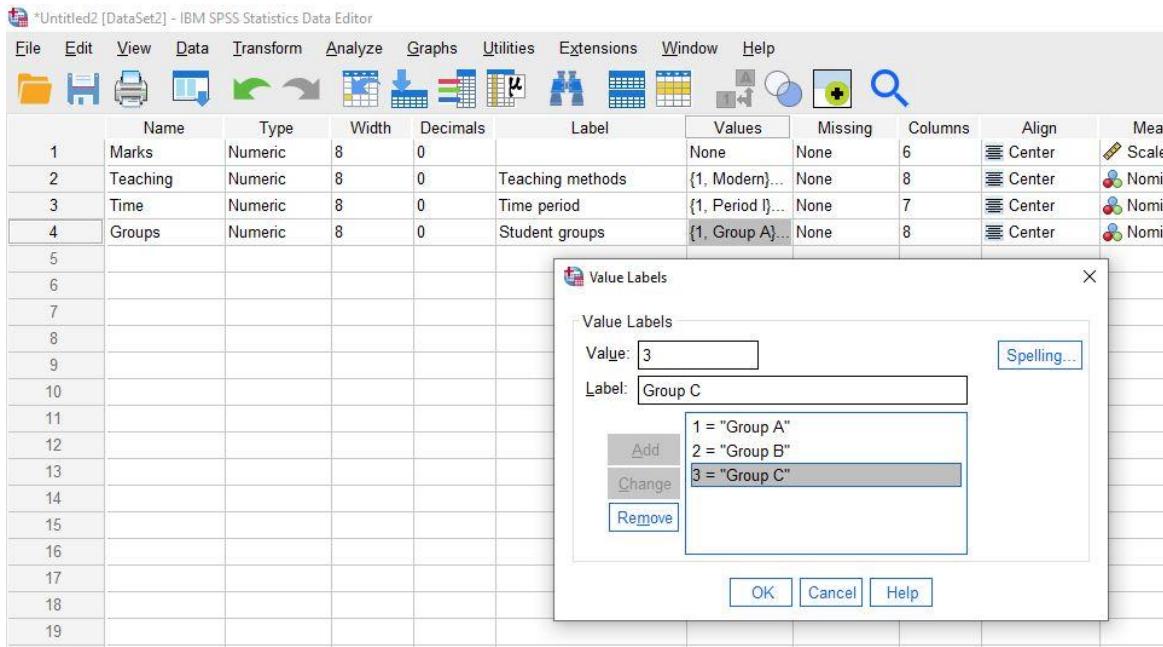
Figure 15.2: Defining variables in Crossover design (SPSS).

Now use the following command sequence:
$$Analyze \rightarrow GLM \rightarrow Univariate$$
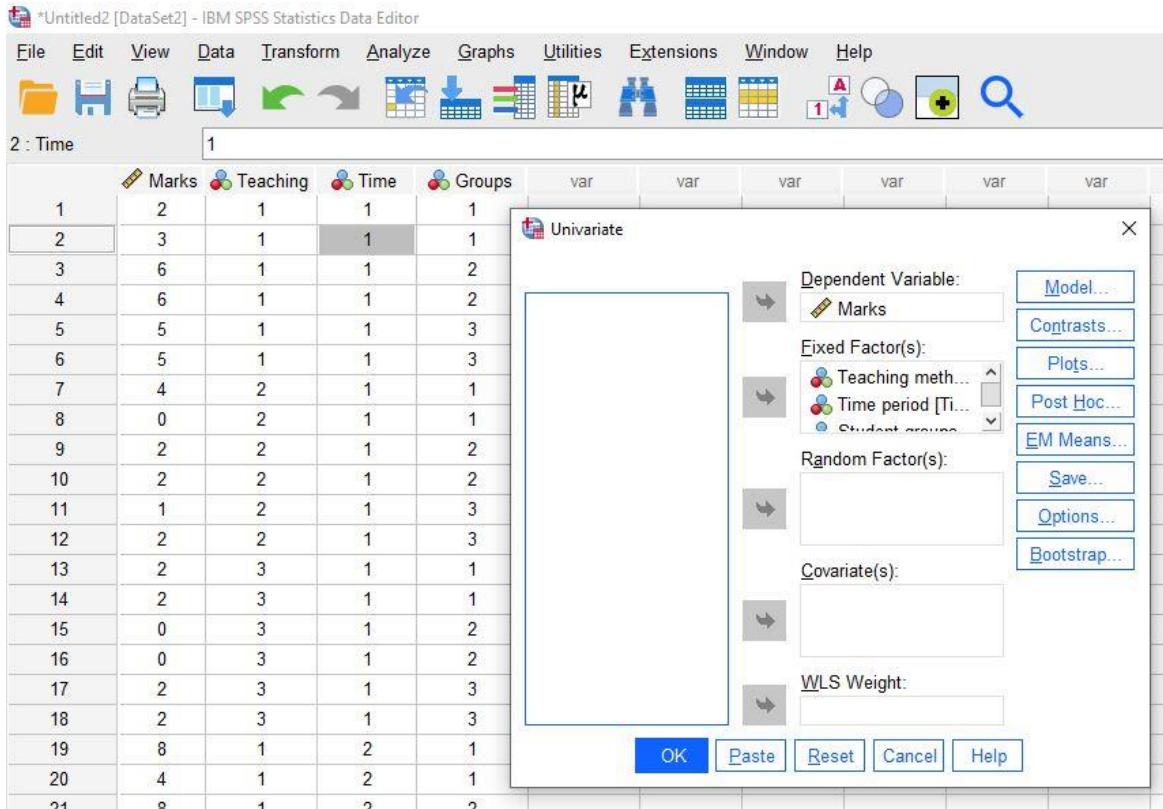
Figure 15.3: Variables input (Crossover design).

All the knowledge belongs to God.

Marks will be **Dependent Variable** and the rest of three factors (Teaching; Time; Groups) as **Fixed Factors** (Figure 15.3). Notice that constraint in execution becomes fixed factor. After selection of **Post Hoc** and **Options** from the tabs, click **OK** to get output:

Figure 15.4: Between-subjects factors key (crossover design).

Table 15.5: Descriptive Statistics output (crossover design)

Descriptive Statistics					
Dependent Variable: Marks					
Teaching methods	Time period	Student groups	Mean	Std. Deviation	N
Modern	Period I	Group A	2.50	.707	2
		Group B	6.00	.000	2
		Group C	5.00	.000	2
		Total	4.50	1.643	6
	Period II	Group A	6.00	2.828	2
		Group B	7.00	1.414	2
		Group C	6.00	.000	2
		Total	6.33	1.506	6
	Total	Group A	4.25	2.630	4
		Group B	6.50	1.000	4
		Group C	5.50	.577	4
		Total	5.42	1.782	12

All the knowledge belongs to God.

Classic	Period I	Group A	2.00	2.828	2
		Group B	2.00	.000	2
		Group C	1.50	.707	2
		Total	1.83	1.329	6
	Period II	Group A	4.00	.000	2
		Group B	2.00	.000	2
		Group C	.00	.000	2
		Total	2.00	1.789	6
	Total	Group A	3.00	2.000	4
		Group B	2.00	.000	4
		Group C	.75	.957	4
		Total	1.92	1.505	12
Mixed	Period I	Group A	2.00	.000	2
		Group B	.00	.000	2
		Group C	2.00	.000	2
		Total	1.33	1.033	6
	Period II	Group A	1.00	1.414	2
		Group B	2.00	.000	2
		Group C	2.00	.000	2
		Total	1.67	.816	6
	Total	Group A	1.50	1.000	4
		Group B	1.00	1.155	4
		Group C	2.00	.000	4
		Total	1.50	.905	12
Total	Period I	Group A	2.17	1.329	6
		Group B	2.67	2.733	6
		Group C	2.83	1.722	6
		Total	2.56	1.917	18
	Period II	Group A	3.67	2.658	6
		Group B	3.67	2.658	6
		Group C	2.67	2.733	6
		Total	3.33	2.567	18
	Total	Group A	2.92	2.151	12
		Group B	3.17	2.623	12
		Group C	2.75	2.179	12
		Total	2.94	2.267	36

All the knowledge belongs to God.
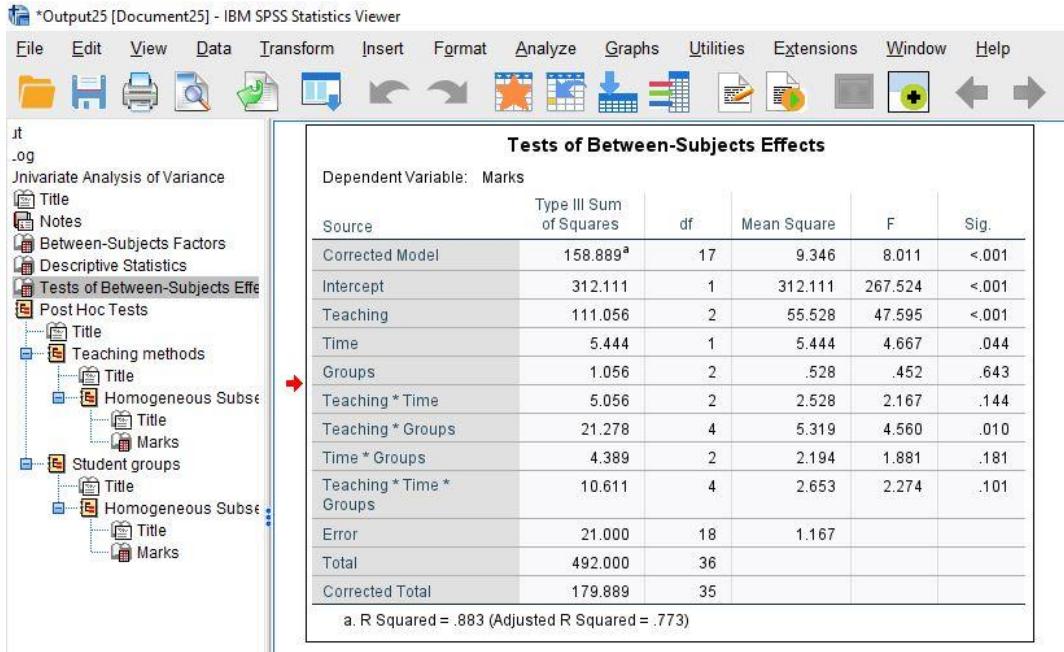
Figure 15.5: Test of Between-Subjects Effect output (crossover design).

As the interaction among **Teaching*Time*Groups** is insignificant ($P = 0.101$), the rest of interactions and main effect become negligible (Figure 15.5). However, if someone is interested in fractional factorial, then **Teaching*Groups** is significant ($P = 0.01$), which means there exist some improvements in marks of students due to Teaching Method's effect. Which can be further elaborated using **Post Hoc** output.
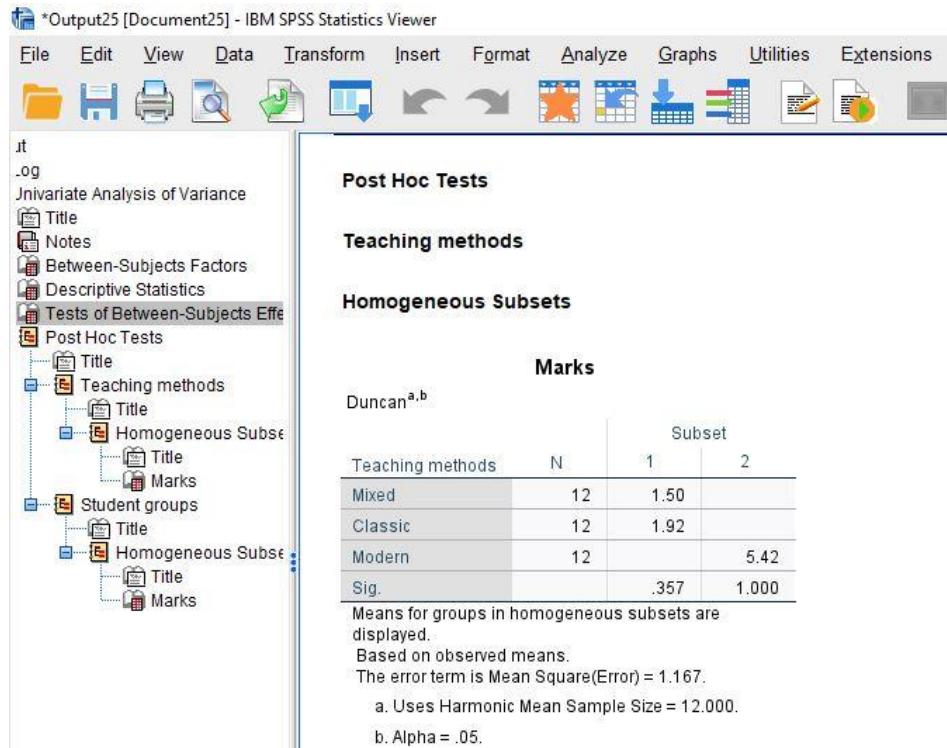
Figure 15.6: Post Hoc for teaching method (crossover design).

All the knowledge belongs to God.

This clearly indicates that the **Modern** method of teaching is significantly different from the other two methods which are (included in subset 1) insignificant for each other (Figure 15.6). Similarly, **Post Hoc** for **Student Groups** can be elaborated as following:
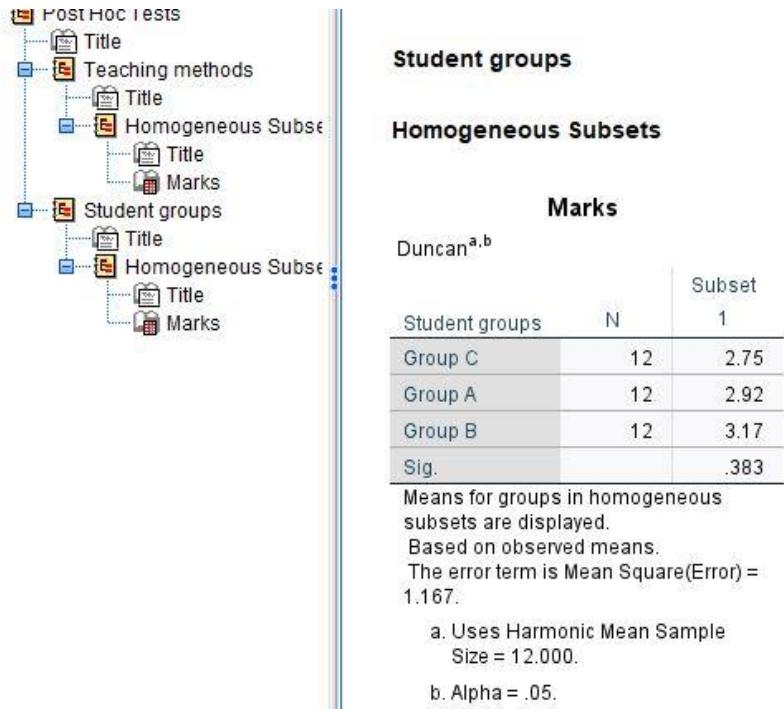
Figure 15.7: Post Hoc for student groups output (crossover design).

As all the groups lie in same subset, these are insignificant for each other.

CHAPTER 16: SPLIT-PLOT DESIGN

All the knowledge belongs to God.

16.1: INTRODUCTION

This design was developed and used first and foremost for agricultural, mainly agronomic experiments, but its applicability goes now across all fields of experimental research. Even so, the terminology for this design still makes references to plots of various types, but the reader should have no difficulty translating this into any other subject matter area. It is unique in its usage. The already planned design is just used in different times in portions. But except for treatment, all other factors are kept same. Other names for such designs, especially those which are nonorthogonal**,** are **bi-randomization designs** and, for the broader class of such structures, **multistratum designs** or **mixed-design ANOVA.**

While accommodating different factors in a single experiment, particularly under field condition, the experimenter may have differential precision requirement for different factors or combinations. Moreover, factors like types of irrigation, types of tillage, types of pest management, drainage management, weed management, etc. require comparatively larger size of plots for convenience compared to the factors like variety, fertilizer, dose, etc. To an experimenter conducting experiment with irrigation and the variety or doses of nitrogen, the varietal effect or the doses of nitrogen or the combination may be more important than the irrigation types; for irrigation treatments the experimenter may need comparatively larger plots than for variety/fertilizer. Thus in such experiments, we are to handle two situations, viz., requirement of differential plot size and differential precision for different treatments. In such cases we opt for split plot designs.

Suppose, if a researcher has planned CR design for 3 types of diets and all the diets are not available to him at the same time. He can only have one diet in one month and other diets after some time, then he need to run CRD for only one treatment level at one time and rest at other time periods. This will be a split-plot as well because researcher has used the design in certain portions. However, the environment is kept same for all treatments levels whether it is run after one month or one year. That is why, it is called **CRD split design** depending on the original design planned.

It is different from the fractional factorial designs where we only have used certain portion of the design and skip the other portions. However, in this design all the portions are used but at different time periods.

16.1.1: Split-Plot Efficiency

Under most circumstances the split-plot design is used for purely technical and practical reasons, as the levels of some factors can be applied only to large EUs which can then be split into smaller EUs for application of the levels of the other factor. This includes also the distinction between **hard-to-change** and **easy-to-change factors** in industrial experimentation.

As an example, consider a feed manufacturer who is interested in three different feed preparation methods (the methods differ in the amount of protein in the feed mixture) and four different pelleting temperatures for the feed and who wishes to study the effect of these two factors on the pellet strength of the feed (Table 16.1). Each replicate of a factorial experiment requires 12 observations, and the experimenter has decided to run three replicates. This will require a total of 36 runs. The experimenter decides to conduct the experiment as follows.

A batch of feed is produced by one of the three methods under study. Then this batch is divided into four samples, and each sample is steamed at one of the four temperatures. Then a second batch of feed is made up using another of the three methods. This second batch is also divided into four samples that are tested at the four

All the knowledge belongs to God.

temperatures. The process is then repeated, until all three replicates (36 runs) of the experiment are obtained. The data is shown in following table:

Table 16.1: The experiment on the tensile strength of paper.

Feed preparation method	Replicate 1			Replicate 2			Replicate 3		
	1	2	3	1	2	3	1	2	3
Temperature (ºC)									
50	30	34	29	28	31	31	31	35	32
60	35	41	26	32	36	30	37	40	34
70	37	38	33	40	42	32	41	39	39
80	36	42	36	41	40	40	40	44	45

Initially, we might consider this to be a factorial experiment with three levels of preparation method (factor A) and four levels of temperature (factor B). If this is the case, then the order of experimentation within each replicate should be completely randomized. That is, we should randomly select a treatment combination (a preparation method and a temperature) and obtain an observation, then we should randomly select another treatment combination and obtain a second observation, and so on, until all 36 observations have been taken.

However, the experimenter did not collect the data this way. He made up a batch of feed and obtained observations for all four temperatures from that batch. Because of the economics of preparing the batches and the size of the batches, this is the only feasible way to run this experiment. A completely randomized factorial experiment would require 36 batches of feed, which is completely unrealistic. The split-plot design requires only 9 batches total. Obviously, the split-plot design has considerable experimental efficiency.

16.1.2: Whole-plot and Split-plot Factors

In this split-plot design we have 9 whole plots (batches in above example), and the preparation methods are called the **whole plot or main treatments**. Each whole plot is divided into four parts called subplots (or split plots), and one temperature is assigned to each. Temperature is so-called the **subplot treatment**.

The terminology of split plots comes from agricultural experimentation, so let's begin with an agricultural example. Suppose that we wish to determine the effects of four corn varieties and three levels of irrigation on yield. Irrigation is accomplished by using sprinklers, and these sprinklers irrigate a large area. Thus, it is logistically difficult to use a design with smallish experimental units, with adjacent units having different levels of irrigation. At the same time, we might want to have small units, because there may be a limit on the total amount of land available for the experiment, or there may be variation in the soils leading us to desire small units grouped in blocks. Split plots give us something of a compromise.

For this purpose, divide the land into six whole plots. These whole plots should be sized so that we can set the irrigation on one whole plot without affecting its neighbors. Randomly assign each irrigation level (represented as I) to two of the whole plots. Irrigation is the **whole-plot (main-plot) factor**, sometimes called the **whole-plot treatment**. Divide each whole plot into four split plots. Randomly assign the four corn varieties (represented as V) to the four split plots, with a separate, independent randomization in each whole plot (Table 16.2). Variety of corn is the **split-plot (subplot) factor**. One Split plots and possible arrangement is as follows, with the six columns representing whole plots with four split plots within each:

All the knowledge belongs to God.

Table 16.2: A typical split-plot layout.

I2 V1	I3 V4	I3 V1	I1 V3	I2 V3	I1 V2
I2 V3	I3 V3	I3 V3	I1 V2	I2 V1	I1 V1
I2 V2	I3 V1	I3 V4	I1 V1	I2 V2	I1 V4
I2 V4	I3 V2	I3 V2	I1 V4	I2 V4	I1 V3

The above terminologies can also be further elaborated. Suppose, each and every replication in split plot design is constituted of number of blocks equal to the number of levels of the factor requiring higher plot size and lesser precision, also known as the **whole plot factor**. Again each and every block should be constituted of as many numbers of homogenous experimental units as the levels of the other factor which require lesser plot size compared to the main plot factor and higher precision, known as the **split-plot factor**. For example, if there be two factors A and B at p and q levels, respectively, included in a split plot experiment and the factor A is the main plot factor, while the factor B is the subplot factor. Similarly, if we have two treatment factors A and B, with levels three. Factor A is referred to as the **whole-plot factor** and the EUs to which the levels of A are applied are the **whole-plots.** Factor B is the **split-plot factor** and the EUs to which the levels of B are applied are the **split-plots.**

There are two separate randomization processes in a split-plot design-one for the main plot and another for the subplot. In each replication, main-plot treatments are first randomly assigned to the main plots followed by a random assignment of the subplot treatments within each main plot.

16.1.3: Split-unit Principle

In all the error-control designs discussed so far, we have had one type of EU for all the treatments and one randomization process to assign the treatments to the EUs. There exist, however, many situations where for a factorial experiment different types of EUs are being used and where the levels of some factors are applied sequentially, necessitating separate randomization procedures. In the simplest case we have EUs of one size for the levels of one of two factors. Those EUs are then subdivided into smaller EUs to which the levels of the second factor are applied. This procedure is referred to as the split-unit principle.

Suppose we want to investigate the pellet strength of poultry feed manufactured using different molasses concentration and steaming it at different temperatures. Let temperature be factor 'a' with three levels, say 50, 70 and 80, and let factor C denote molasses concentration with four levels (Table 16.3). We have three pellet machines available. Each machine will be set at one of the randomly assigned temperatures. In each machine we then place four pellet feed bags each individually produced using a different (randomly assigned) molasses concentration for each bag of pellet. This process is repeated on several days. For each pellet bag the breaking strength is then determined using a suitable machine. The large EUs are pellet machines, and the smaller EUs are the pellet feed bags:

Table 16.3: Breaking strength of plates.

	a3		**a1**		**a2**	
DAY 1	C3	C4	C3	C4	C3	C4
	C2	C1	C2	C1	C2	C1
	a1		**a3**		**a2**	
DAY 2	C3	C4	C3	C4	C3	C4
	C2	C1	C2	C1	C2	C1

All the knowledge belongs to God.

Not only will this type of arrangement lead to different precisions for the comparisons among the levels of the a-factor and among those of the C-factor, but the fact that the two factors are associated with different types of EUs leads to different experimental error variances associated with these comparisons. This is the reason why these types of experiments must be distinguished very carefully from the factorial experiments.

In order to better understand the structure of the simple split-plot design it is advantageous to view it as superimposing one RCBD (for the split-plot treatments) on top of another RCBD (for the whole-plot treatments) known as a **superimposed plot design** (RCBD, RCBD). Variations of this form of split-plot design are possible by using different component designs, other than both RCBD.

16.2: PROS AND CONS OF SPLIT-PLOT DESIGN

Managing different factors as per the requirement without sacrificing the information from the design is the main advantage. Different factor effects are estimated at different precision levels as well which was not possible in simple experiments. So, in split plot design the subplot factor and its interaction with the main plot factor are estimated more precise than the main plot factor effect.

However, the factor to be assigned as the main plot factor or subplot factor is of extremely importance. Randomization and layout along with the analysis are also somewhat complicated compared to other simple experiments. The comparison of the main plot treatment means at the same or different levels of the sub treatment is somewhat approximately 1. Furthermore, when both the factors require larger plot size then split plot design is not suitable.

16.3: EXAMPLE

16.3.1: Problem

To know the effect of three different feeds and four doses of choline on the weight of broiler chicken an experiment was conducted (Table 16.4). The experiment was conducted in nine poultry houses comprising of three blocks each having three pens. Each pen was then partitioned into four big cages. Three different feeds were randomly allocated to each of the three houses separately in each pen. Among the four cages of each house, four dose of choline were allocated randomly. Body weights (kg) are reached at the age of 35 days. Analyze the data and draw your conclusion:

Table 16.4: Weight of broiler chicken.

	Feed 1				Feed 2				Feed 3			
Choline	**T1**	**T2**	**T3**	**T4**	**T1**	**T2**	**T3**	**T4**	**T1**	**T2**	**T3**	**T4**
Block 1	1.83	1.88	1.92	2.01	1.85	1.91	1.94	2.05	1.86	1.93	1.97	2.22
Block 2	1.84	1.89	1.93	2.03	1.86	1.92	1.95	2.09	1.86	1.93	1.96	2.17
Block 3	1.81	1.88	1.93	2.05	1.87	1.91	1.95	1.95	1.85	1.94	1.97	2.19

16.3.2: Solution

There are following variables included:
- $Factor\ 1 = feed\ (3\ levels)$
- $Factor\ 2 = choline\ (4\ levels\ as\ T1; T2; T3; T4)$
- $Constraint = block\ (3\ levels)$

The hypotheses will be a little bit different. Here in this experiment each poultry house in each block can be treated as the main plot in this 'split plot design' and four cages in each pen as the subplots. Thus, we have **three main plots** and **four subplots.** Main plot and subplot factors are feed (F1, F2, and F3) and doses of choline (T1, T2, T3 and T4), respectively. The experiment is repeated three times, i.e., in three blocks.

For the above experiment, we have three levels of main plot factors and four levels of subplot factors in three replications. That is why there would be three hypotheses in it.

All the knowledge belongs to God.

$$H_1 = Change\ in\ feed\ has\ an\ effect\ on\ chick\ weight$$

$$H_1' = Change\ in\ trilostane\ has\ an\ effect\ on\ chick\ weight$$

$$H_1'' = Change\ in\ block\ has\ an\ effect\ on\ chick\ weight$$

The SPSS procedure will depend on the basic design which has been split and the coding used. If someone knows the writing coding in SPSS, it will be very easy to analyze the data. For example, if the CRD is used as split-design then the one-way ANOVA will be executed in SPSS. In this case, it will be the same as 3-way ANOVA (for LSD).

All the knowledge belongs to God.

CHAPTER 17: NESTED DESIGN

All the knowledge belongs to God.

17.1: INTRODUCTION

Nested designs are usually the result of shortage of subjects or some other limitation on experimental units. A nested design is one that uses replication of experimental units in at least two levels of hierarchy. That is why it is also known as **hierarchical designs.**

These designs are a way to use replication to increase one's confidence that differences seen during treatment comparison are real and not just by chance. Because these designs result in data from replicated samples taken from replicated plots receiving each treatment of interest. They are normally used when you want to include more than one task under a level of an independent variable. We actually create sub-blocks to do these experiments.

It is a design, where response for each factor is perceived differently by the subjects under study i.e., if we have 3 people from Lahore, Gilgit and Karachi and we want to check the effect of temperature of Murree on their health. Each subject will perceive the Murree's temperature differently. Murree's feel temperature will be much colder for Karachi's guy as compared to the Gilgit's guy. It is the factor (feel temperature) under a factor (temperature). That is why it is known as Nested-design, just like a mother protects her babies under herself. However, both factors (temperature and feel temperature) are not the same, otherwise it will be a factorial design. These designs are more economical than factorial designs but they do not yield as much information as factorial.

17.1.1: Simple Nested Design

Consider a company that purchases its raw material from three different suppliers (Figure 17.1). The company wishes to determine whether the purity of the raw material is the same from each supplier. There are four batches of raw material available from each supplier, and three determinations of purity are to be taken from each batch.
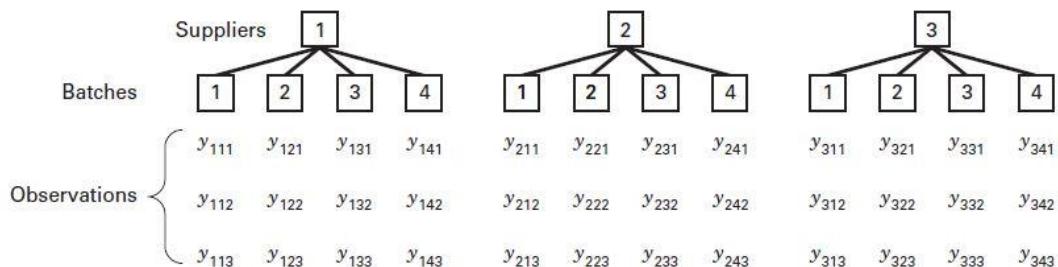
Figure 17.1: Two staged nested design layout.

This is a two-stage nested design, with batches nested under suppliers. At first glance, you may ask why this is not a factorial experiment. If this were a factorial, then batch 1 would always refer to the same batch, batch 2 would always refer to the same batch, and so on. This is clearly not the case because the batches from each supplier are unique for that particular supplier. That is, batch 1 from supplier 1 has no connection with batch 1 from any other supplier, batch 2 from supplier 1 has no connection with batch 2 from any other supplier, and so forth. To emphasize the fact that the batches from each supplier are different batches, we may renumber the batches as 1, 2, 3, and 4 from supplier 1; 5, 6, 7, and 8 from supplier 2; and 9, 10, 11, and 12 from supplier 3, as shown in above figure.

All the knowledge belongs to God.

Take another example (Figure 17.2) which includes three levels of between-subjects factor (A1, A2 and A3). Under each of the three levels of factor A are nested three levels of factor B (B1, B2 and B3).
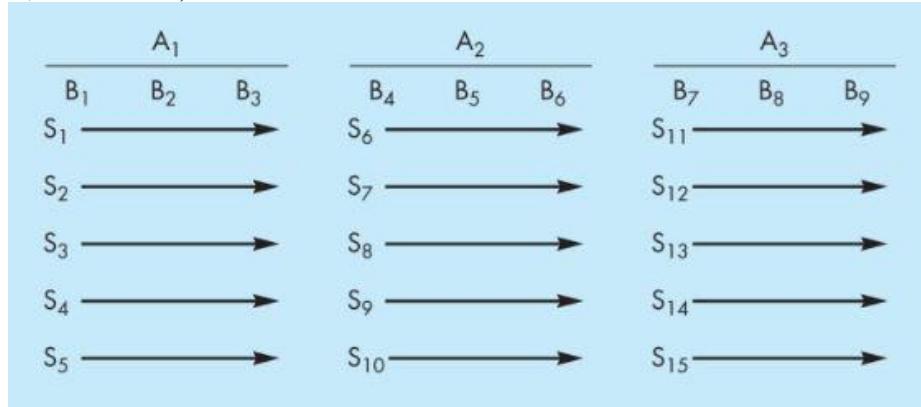
Figure 17.2: Three-stage nested design layout.

Each factor A level thus includes within-subjects factor. Note that we have combined within-subjects factors and between-subjects factors in nested design.

Occasionally in a multifactor experiment, some factors are arranged in a factorial layout and other factors are nested. We sometimes call these designs **nested–factorial designs**.

17.2: EXAMPLE

17.2.1: Problem

Suppose we were fancy farmers concerned with keeping our rare chicken (Silkie) free of lice. We wish to find out whether a topical application of a quick-acting pyrethrin insecticide reduces the number of lice per feather in males, females and young chickens. Our fancy farm has only one pair of Silkie and their single young. This shortage may force us for nested design. We apply a dose of pyrethrins topically to one area on each bird and use another area of the same bird as the control. Thus, pyrethrins and the control treatments are nested within a bird.

17.2.2: Layout

Here we have K replicates of three Silkie types and we apply control and pyrethrins on birds as treatments (Figure 17.3). So, we compare the differences in between the two treatments within each bird only, rather among a random sample of chickens.
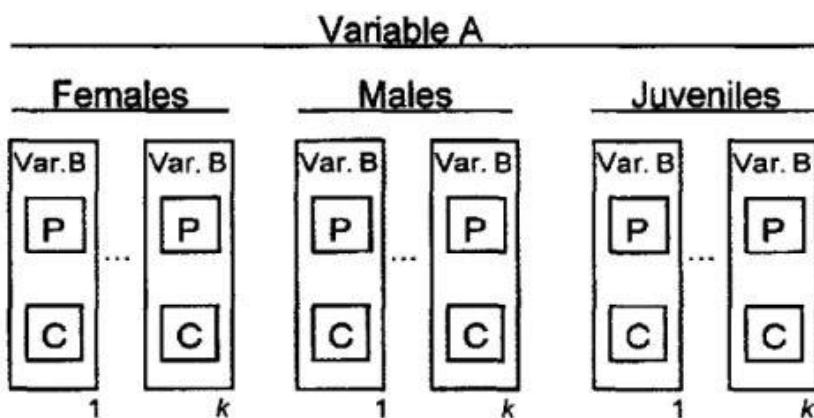
Figure 17.3: Pyrethrin application on cassowaries.

All the knowledge belongs to God.

CHAPTER 18: KEY OF DESIGN OF EXPERIMENT

Table 18.1: Key features of different designs in DOE.

Design	Response	EUs[1]	Constraint Presence	(#)	(#)	Factor Sub-factor	No. of Runs	Factorial	ANOVA	Treatments (Runs)	(Repetition)	WSF[2]
CRD[3]	NN[4]	Hom.[5]	No	0	1	Nil	Single	NO[6]	1-way	All at one time	No	No
RCBD[7]	NN	Het.[8]	May be	1	1-2	Nil	Single	May be	2-way	All at one time	No	No
LSD[9]	NN	Het.	May be	2	1-3	Nil	Single	May be	3-way	All at one time	No	No
GLSD[10]	NN	Het.	May be	3	1-4	Nil	Single	May be	4-way	All at one time	No	No
SPD[11]	NN	Any	May be	Any	1-poly	Nil	Single	May be	Any	At different times	NO	No
ND[12]	NN	Hom.	May be	Any	1-poly	YES	Single	May be	Any	All at one time	No	No
COD[13]	NN	Any	May be	Any	1-poly	Nil	>1	May be	Any	At different times	NO	YES
RMA[14]	NN	Any	May be	Any	1-poly	Nil	>1	May be	RMA	At different times	YES	YES
ANCOVA	NN	Any	COV[15]	Any	1-poly	Nil	Single	May be	ANCOVA	All at one time	No	NO
None	Nominal	No need to look at any other feature										

Above table contains the common features of different designs used in experiments. The most unique feature of a particular design is highlighted in **yellow** which can be helpful for a single guess of identification.

---

[1] Experimental units
[2] Within-Subjects Factor
[3] Completely randomized design
[4] Non-nominal
[5] Homogenous
[6] Factorial CRD is possible only in case when two-way ANOVA is converted into one-way ANOVA. But not in its pure form.
[7] Randomized complete block design
[8] Heterogeneous
[9] Latin square design
[10] Graeco-Latin square design
[11] Split-plot design
[12] Nested design
[13] Crossover design
[14] Repeated Measure ANOVA
[15] Covariance

All the knowledge belongs to God.

CHAPTER 19: NON-PARAMETRIC TESTS IN SPSS

19.1: INTRODUCTION

Nonparametric tests are used in a situation when data is measured on nonmetric scale. In other words, if the measurements are categorical in nature, these tests are used in hypothesis testing experiment. Normally these tests are used because the dependent variable is not interval- or ratio-scaled. Nonparametric tests can also be used for metric data if assumptions of parametric tests are severely violated.

In research studies, quite often distribution of the population from which the sample is drawn is unknown, and therefore in such situations nonparametric tests are the best option. In using nonparametric tests, no assumption is made about the distribution of the population from which the samples are obtained; hence, these tests are also known as **distribution-free test** or **assumption-free tests**, with an explanation that they make no assumptions about the distribution of the data. Technically, this isn't true: they do make distributional assumptions (e.g., the ones in this chapter, all assume a continuous distribution), but they are less restrictive ones than their parametric counterparts.

Not many assumptions are required for using nonparametric tests; hence, it can be easily used by the researchers. On the other hand, all parametric tests are based on the assumption that the distribution of the population from which the samples are drawn is normal. Besides this, each parametric test requires certain assumptions to be made. Thus, in all those situations where normality assumption is violated, required assumptions of the parametric test breaks down, or if the data is measured on nominal or ordinal scales, nonparametric tests are used in hypothesis testing.

19.1.1: Parametric versus Non-Parametric Tests

Plenty of differences exists between two. In parametric tests, hypothesis concerning proportion, mean, or variance is usually tested, whereas in nonparametric tests, hypothesis concerning median is tested for its significance. In parametric tests, parent population is assumed to be normally distributed.

All tests that we have read so far like z, t, and F are known as parametric tests. Since these tests investigate the hypothesis concerning parameters, they are known as parametric tests. On the other hand in nonparametric tests distribution of the population need not to be normally distributed. Since hypothesis testing does not involve any parameter, these tests are known as nonparametric tests. For example, assessment of playing ability, quality of cricket shot, and performance in soccer produce data on ordinal scale. On the other hand performance on minimum muscular fitness test and match result in hockey results in nominal data as these performances are assessed on pass/fail or winning/loosing format, respectively. In all such situations, nonparametric tests are well suited for hypothesis testing.

The procedure of testing hypothesis is similar in nonparametric and parametric tests. The only difference is in terms of constructing hypothesis and computing test statistic. In most of the commonly used parametric tests, there is an alternative test available in nonparametric. For example, **Mann–Whitney U tests** are alternative tests to the two-samples t-test, and **Wilcoxon signed-rank test** is an alternative to the paired t-test. Similarly, nonparametric tests such as **Kruskal–Wallis** and **Friedman** can be used as an alternative to the one-way ANOVA and repeated measures ANOVA, respectively.

Strictly speaking, only those procedures that test hypotheses that are not statements about population parameters are classified as nonparametric, while those that make no assumption about the sampled population are called distribution-free procedures. Despite this distinction, it is

customary to use the terms nonparametric and distribution-free interchangeably and to discuss the various procedures of both types under the heading nonparametric statistics.

19.1.2: Pros and Cons of Non-Parametric Tests

The above discussion implies the several advantages of nonparametric statistics. They allow for the testing of hypotheses that are not statements about population parameter values. Some of the chi-square tests of goodness-of-fit and the tests of independence are examples of tests possessing this advantage. Nonparametric tests may be used when the form of the sampled population is unknown.

Nonparametric procedures tend to be computationally easier and consequently more quickly applied than parametric procedures. This can be a desirable feature in certain cases, but when time is not at a premium, it merits a low priority as a criterion for choosing a nonparametric test. Indeed, most statistical software packages now include a wide variety of nonparametric analysis options, making considerations about computation speed unnecessary. Similarly, nonparametric procedures may be applied when the data being analyzed consist merely of rankings or classifications. That is, the data may not be based on a measurement scale strong enough to allow the arithmetic operations necessary for carrying out parametric procedures. The subject of measurement scales is discussed in more detail in the next section.

Although nonparametric statistics enjoy a number of advantages, their disadvantages must also be recognized. The use of nonparametric procedures with data that can be handled with a parametric procedure result in a waste of data. The application of some of the nonparametric tests may be laborious for large samples.

19.2: KOLMOGOROV–SMIRNOV ONE-SAMPLE TEST.

This test is the alternative to one-sample t-test in non-parametric version.

19.2.1: SPSS Procedure

Experimental Design marks (out of 10) of 18 PhD students are shown in Table 19.1. We wish to know if we may conclude that these data are not from a normally distributed population with a mean of 8 and standard deviation of 2.

Table 19.1: Marks of PhD students.

8	8	7
6	4	8
4	5	0
3	6	5
0	7	6
4	10	6

For analysis, it will be the same as One-sample T Test (parametric version), but as the normality assumption has been violated, the test will be different. Variables included in above example are:

- Dependent variable=marks
- Test value=8

The alternative hypothesis will be:

$$H_1 = The\ mean\ marks\ of\ students\ is\ not\ 8$$

All the knowledge belongs to God.

Enter data into **Data View** module and follow the commands as:

$$Analyze \rightarrow Nonparametric\ Tests \rightarrow Legacy\ Dialogue \rightarrow 1-Sample\ K-S$$
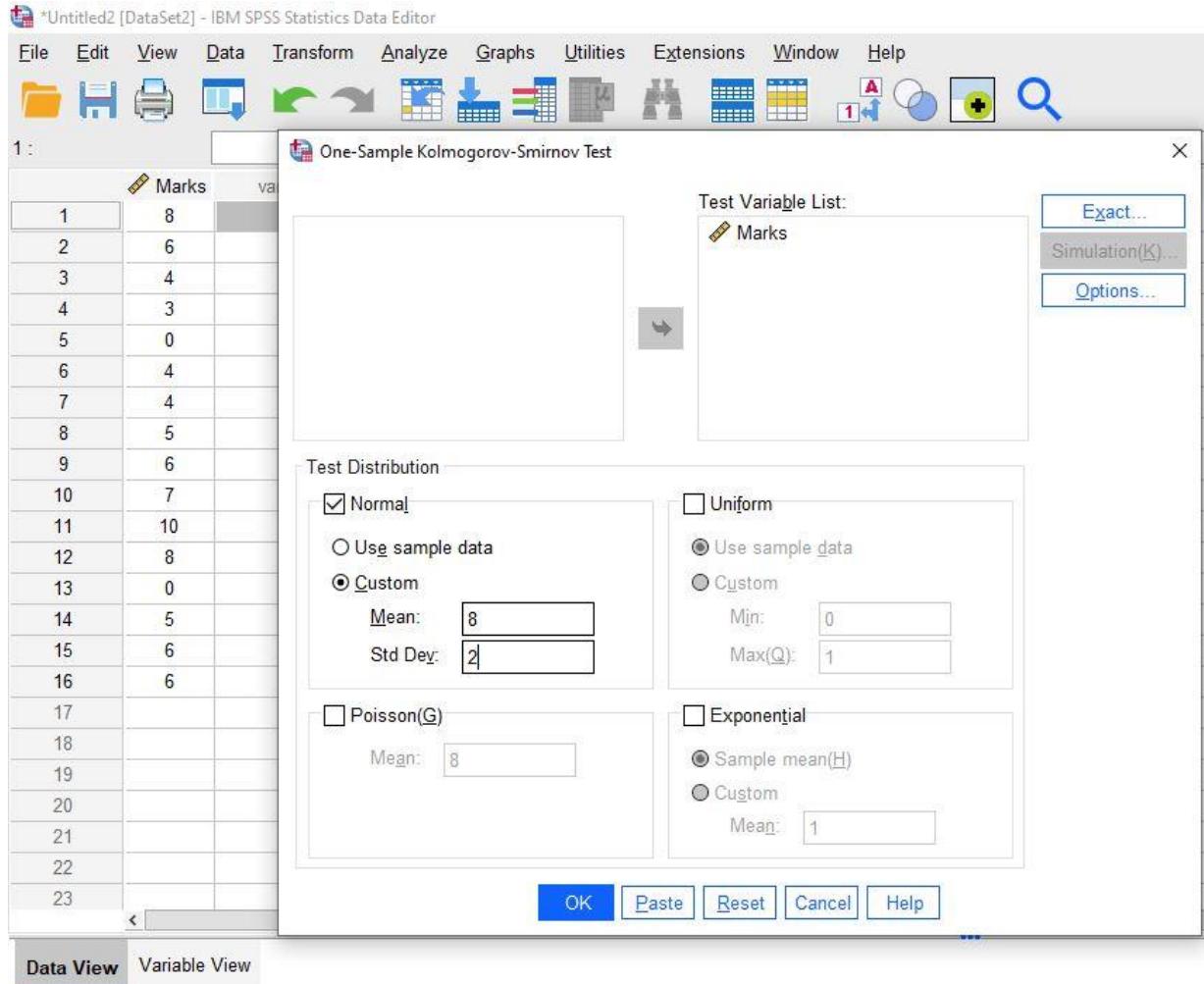
Figure 19.1: Variable input (1-Sample K-S).

Select the **Custom** option under **Test Distribution** and put the value 8 (test value) in **Mean** option (Figure 19.1). Select the **Options** and click continue. Click **OK** to get output.
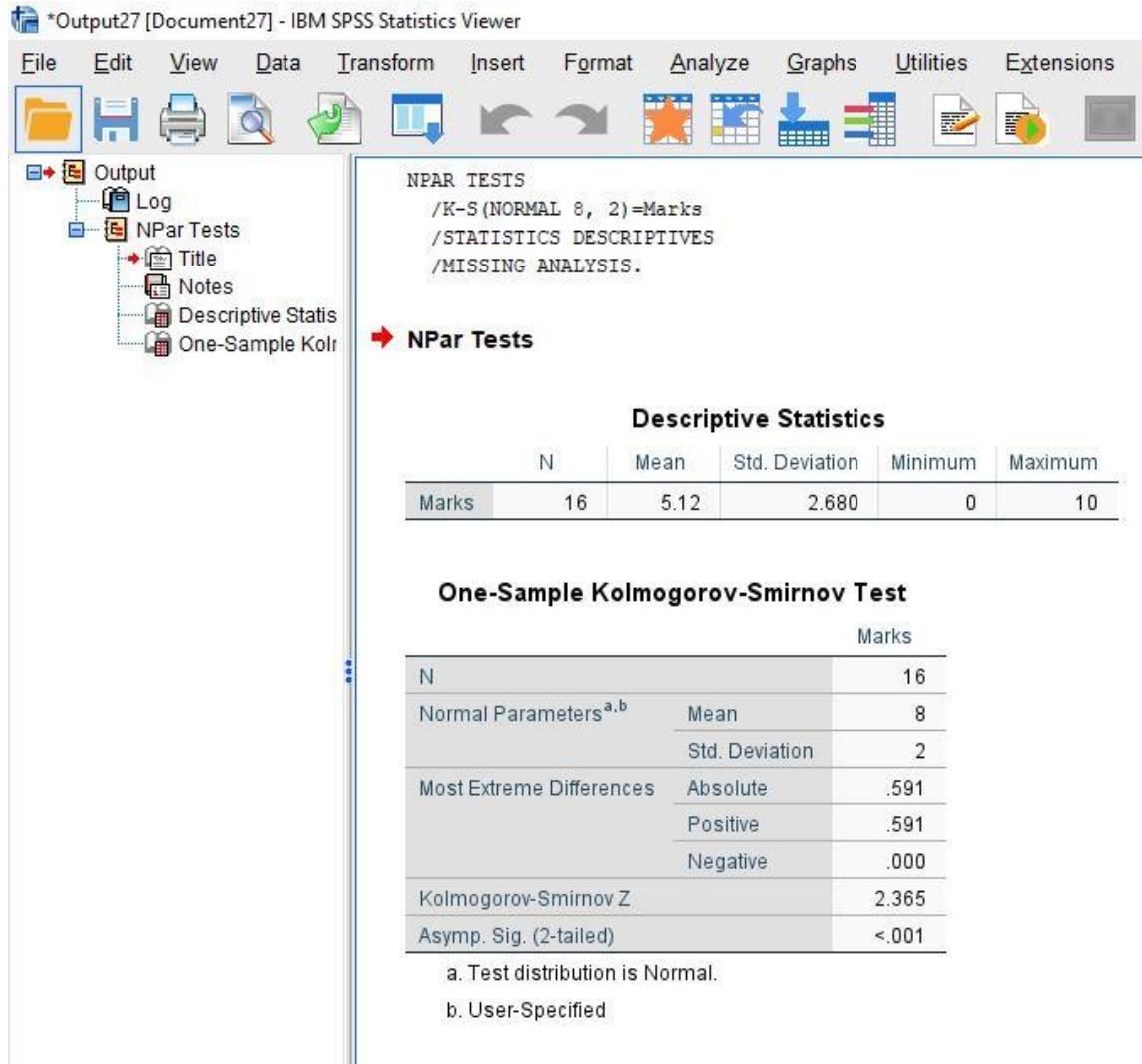
Figure 19.2: Descriptive Statistics and One-Sample KS Test output (1-Sample K-S).

The '**Asymp.Sig**' is significant (Figure 19.2) which means H_1 is accepted.

19.3: MANN–WHITNEY U TESTS

When you want to test differences between two conditions and different participants have been used in each condition then you have two choices: the **Mann–Whitney test** and the **Wilcoxon rank-sum test**. These tests are the non-parametric equivalent of the independent t-test. In fact, both tests are equivalent, and there's another, more famous, Wilcoxon test.

Mann-Whitney is alternative to two-independent samples test. If assumptions of t-test are seriously violated, then the Mann–Whitney U test (sometimes called the Mann–Whitney–Wilcoxon test) may be used to compare two independent groups. In this test, no assumption is required about the distribution of population from which the samples have been drawn. In using Mann–Whitney U test we intend to test whether the two samples come from the same population or not. Since this test can be used both for parametric and nonparametric data, it is the most powerful nonparametric test. The Mann–Whitney U test can be used more efficiently as an alternative to the t-test if we wish to avoid the assumptions like equality of variance and normality of the population distribution. The

All the knowledge belongs to God.

only assumption in using this test is that the data must be measured at least on the ordinal scale and samples must be randomly drawn.

19.3.1: SPSS Procedure

DVM students are considered to be less active than Poultry students due to the nature of their curriculum. To investigate this fact, a study was planned in which 12 students were randomly chosen from the DVM as well as from the Poultry classes. These students were tested for physical test score (out of 100) specially designed by a physical trainer. The data so obtained are shown in Table 19.2. The assumptions of t-test are seriously violated; hence, let us investigate the research question.

Table 19.2: Physical test score of students.

DVM	63	60	62	64	60	63	54	55	59	65	59	69
Poultry	70	66	58	65	62	55	77	57	78	80	55	61

The test variables include:
- $Test\ Variable\ List = test\ score$
- $Grouping\ Variable = degree$

We need to test the following hypothesis:

$$H_1 = Physical\ test\ score\ is\ not\ similar\ in\ both\ the\ groups$$

The Mann-Whitney U test uses the rankings of the data. Therefore, the data for the two samples must at least be ordinal. To prepare data file, define Degree as a nominal and Test Score as scale variables in the **Variable View**. For Degree variable, give code 1 to DVM and code 2 to Poultry by clicking the cell under the column heading **Values.**

After defining variables, the screen shall look like as shown in the following figure:
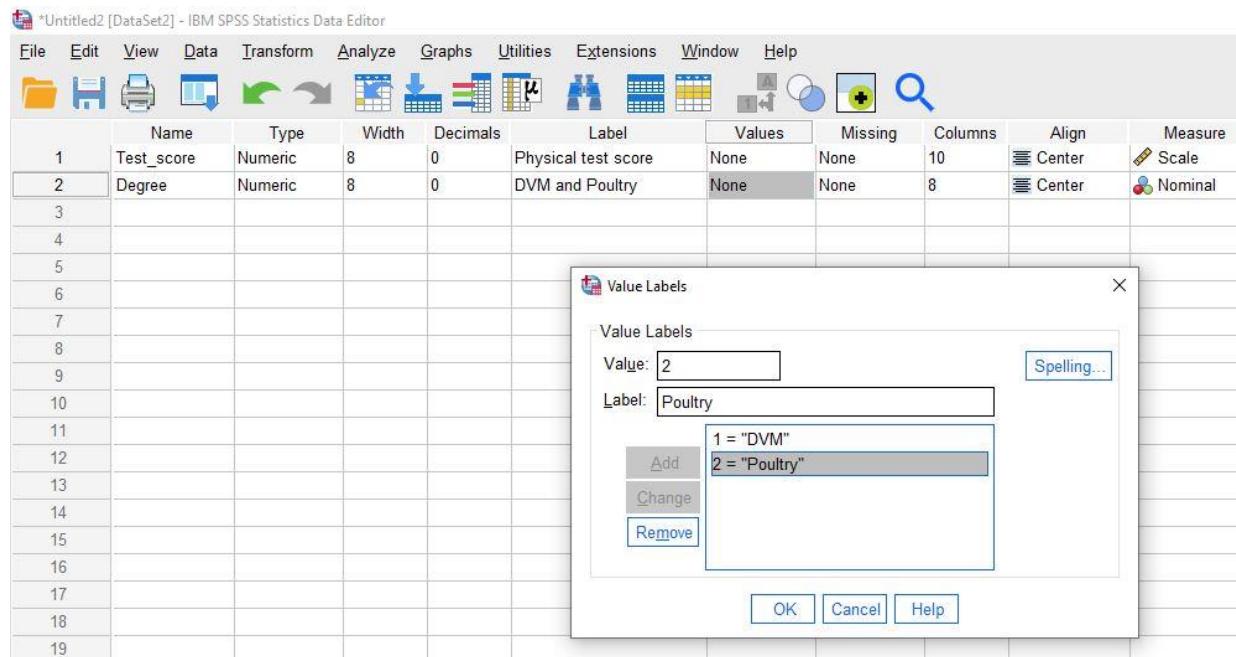
Figure 19.3: Defining variables (Mann-Whitney U test).

All the knowledge belongs to God.

After entering data in the **Data View**, do the following steps:
$$Analyze \rightarrow Nonparametric\ Tests \rightarrow Legacy\ Dialogs$$
$$\rightarrow 2\ Independent\ Samples\ ....$$
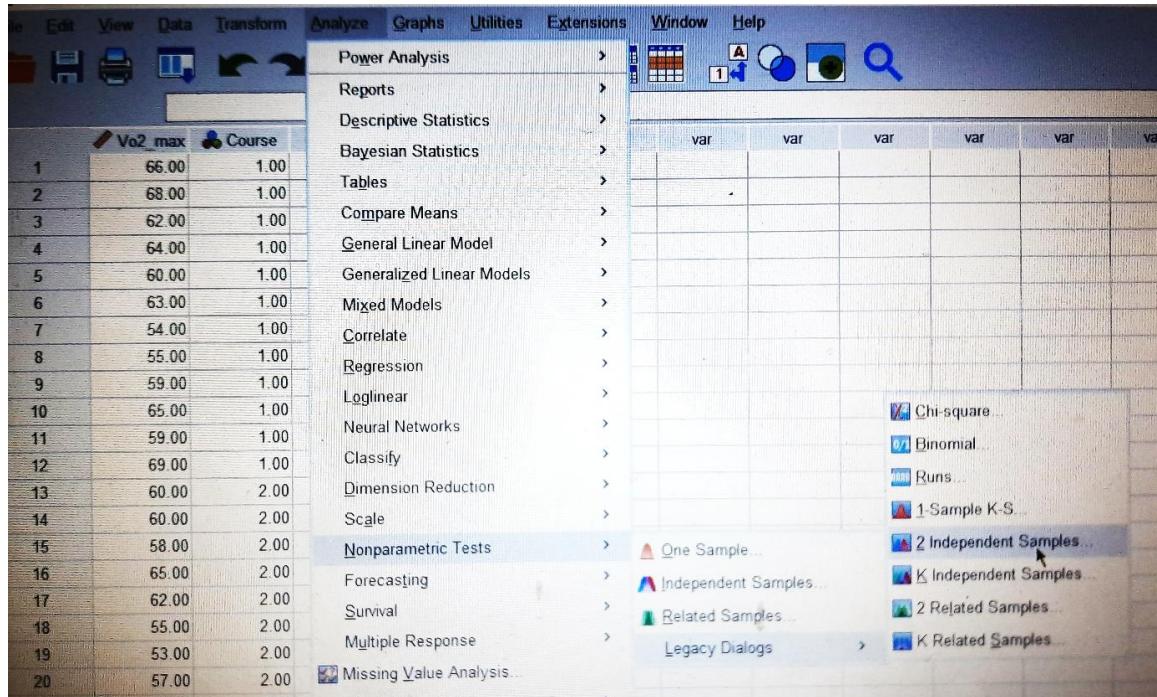
Figure 19.4: Sequence of command for Mann-Whitney U test (SPSS).

Do the same as previously explained parametric test (Independent Samples T Test) and select options (Figure 19.5). Select Test_score and Degree variables from the left panel and bring them into the **Test Variable List** and **Grouping Variable** sections in the right panel.
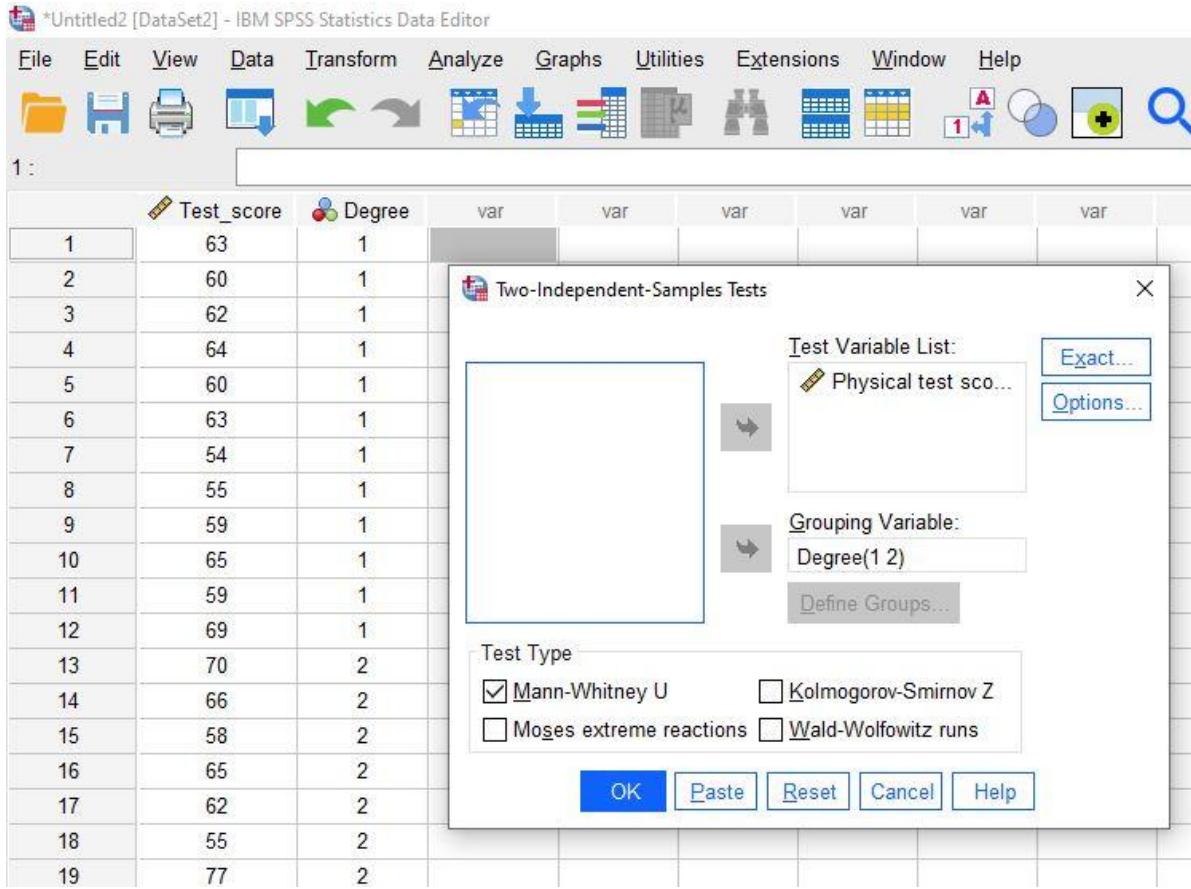
Figure 19.5: Variable input (Mann-Whitney U Test).

Enter 1 and 2 in **Group 1** and **Group 2**, respectively, as shown in the Figure 19.6. Ensure that the **Mann–Whitney U** option is checked.
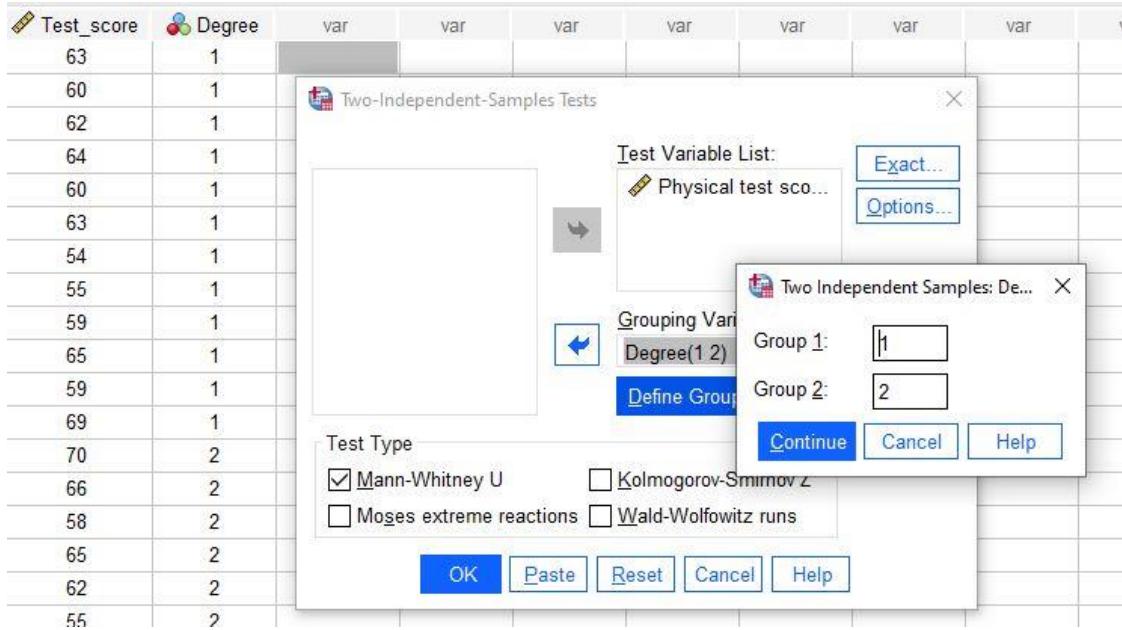
Figure 19.6: Define Groups editing (Mann-Whitney Test).

All the knowledge belongs to God.

Click on **Options** command and check **Descriptive** as following figure:
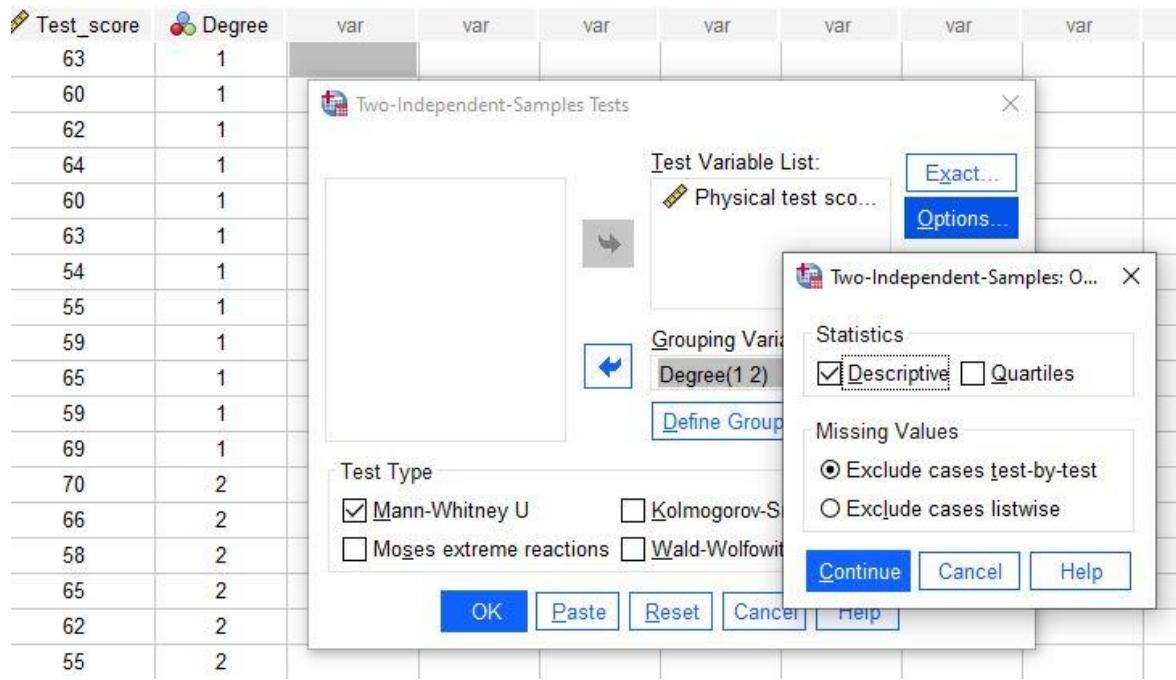
Figure 19.7: Options selection (Mann Whitney Test).

Click on **Continue** and **OK** options to get the outputs:
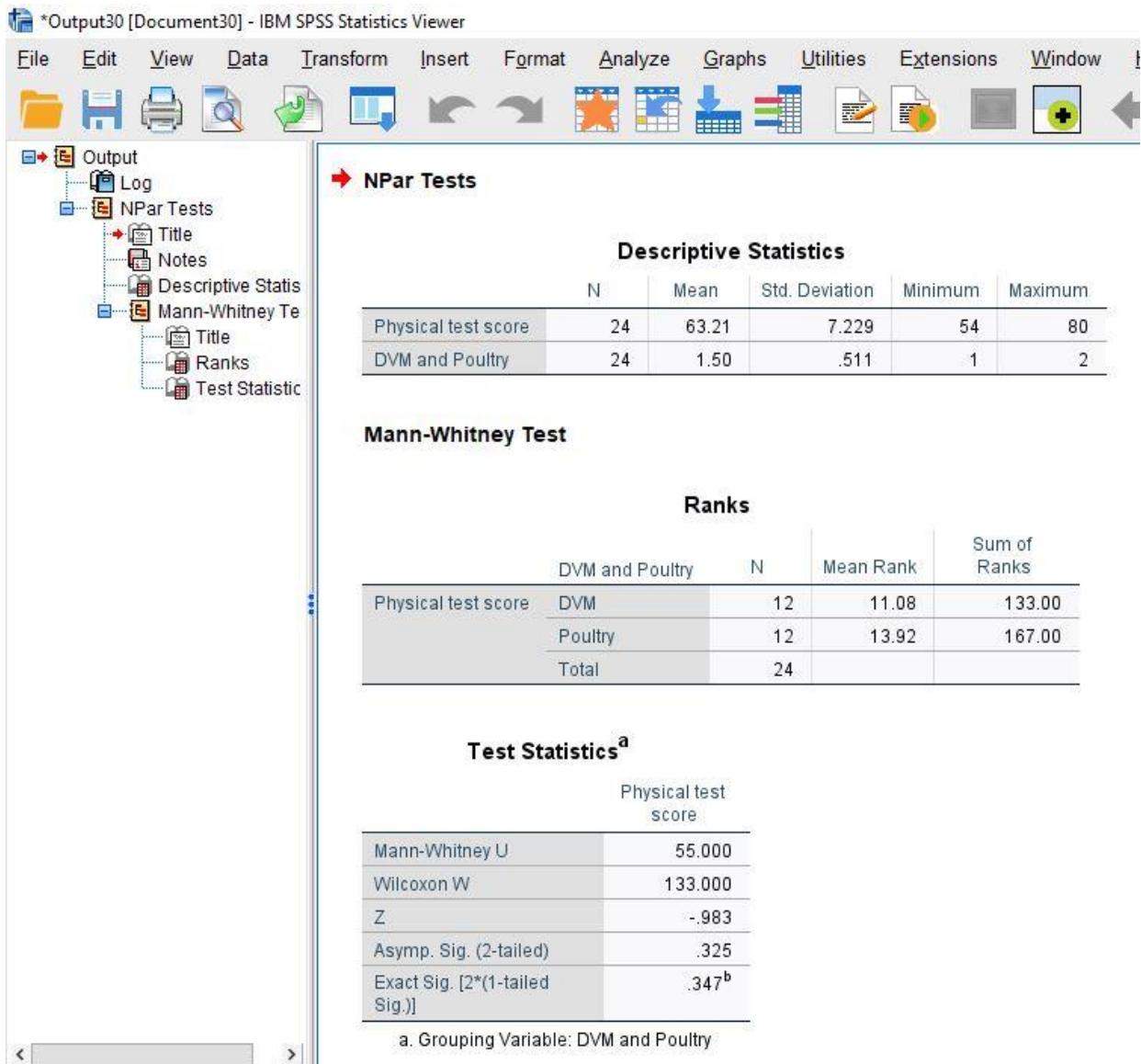
Figure 19.8: Descriptive Statistics and Ranks output (Mann-Whitney Test).

Above Figure 19.8 shows the mean rank in both the groups, the Mann–Whitney U statistic and other results. The Mann–Whitney test works by looking at differences in the ranked positions of scores in different groups. Therefore, the first part of the output summarizes the data after they have been ranked. Specifically, SPSS tells us the average and total ranks in each condition. Remember that the Mann–Whitney test relies on scores being ranked from lowest to highest; therefore, the group with the lowest mean rank is the group with the greatest number of lower scores in it. Similarly, the group that has the highest mean rank should have a greater number of high scores within it. Therefore, this initial table can be used to ascertain which group had the highest scores, which is useful in case we need to interpret a significant result.

The second table provides the actual test statistics for the Mann–Whitney test, the Wilcoxon procedure and the corresponding z-score Figure 19.8. To interpret the findings, we need to consider the z statistic and two tailed p value, corrected for ties (**Asump.Sig** value). The output

indicates that the **Asymp.Sig** is insignificant (0.325), p > 0.05; hence, significant differences in Physical Test Score does not exist between DVM and Poultry students.

The **Asymp. Sig.** (asymptotic significance) means this is not an exact significance level (in comparison to **Exact Sig**. The p < 0.325 indicates that there is no statistically significant overall difference among the mean ranks.

19.4: WILCOXON SIGNED-RANK TEST

The Wilcoxon signed-rank test, not to be confused with the Wilcoxon rank-sum test in the previous section, is used in situations in which there are two sets of scores to compare, but these scores come from the same participants (when the data in both groups is related in some sense and we wish to test whether the members of a pair differ). The Wilcoxon Signed-Rank test is a nonparametric alternative to the paired t-test. The Wilcoxon signed-rank test does not require that the data be a sample from a normally distributed population.

Figure 19.9: Frank Wilcoxon.

19.4.1: SPSS Procedure

A gym training program was launched by a university in which 8 female PhDs students participated. Their weights were measured before and after the program. The data obtained in the study is shown in Table 19.3. The data violates the assumption of paired t-test. Let us investigate whether the program was effective in reducing participant's weight at 5% level.

Table 19.3: Student's weight.

Pre-program	75	80	89	92	74	72	70	68
Post-program	69	79	90	88	68	70	69	68

Variables included in above problem are:
- Test pairs=Variable 1 (pre prog) and variable 2 (post prog)
- Response=student's weight

We need to test the following hypothesis:

$H_1 = Significant\ difference\ exists\ between\ pre\ and\ post\ program\ weight\ loss$

All the knowledge belongs to God.

The Wilcoxon test is based on the difference in rankings. The data for the two samples must at least be ordinal. Before applying Wilcoxon Signed-Rank test, a data file needs to be prepared. This can be done by defining the variables Pre_prog and Post_prog as **Scale** in **Variable View** (Figure 19.10). Under the heading **Label,** expanded name of the variables can be defined.

Figure 19.10: Defining variable (Wilcoxon Signed-Rank Test).

After defining variables, click on **Data View** to open the format for entering the data column-wise. Notice the difference of data feeding layout which is different than the one we used in case of Mann–Whitney U test. In **Data View**, click on the following commands in sequence:
Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples

After clicking **2 Related Samples** option, you will be taken to the screen as shown in the Figure 19.11 for selecting variables and selecting option. Select Pre_prog and Post_prog variables from the left panel and bring them into the **Test Pairs** section in the right panel. Ensure that the **Wilcoxon** option is checked.

Figure 19.11: Variables input (Wilcoxon Signed Rank test).

All the knowledge belongs to God.

To carry out several Wilcoxon tests select another pair of variables, transfer them to the variables list, and then select another pair and so on. Each pair appears as a new row in the box labelled **Test Pairs**. Click on **Options** command and check **Descriptive**.

If you click on **Exact** then another dialog box appears (Figure 19.12) that allows you to select for SPSS to compute exact significance values. When samples are large you should probably opt for the **Monte Carlo** method, and when you have small samples it's worth opting for the **Exact** test. I haven't opted for either in this example.
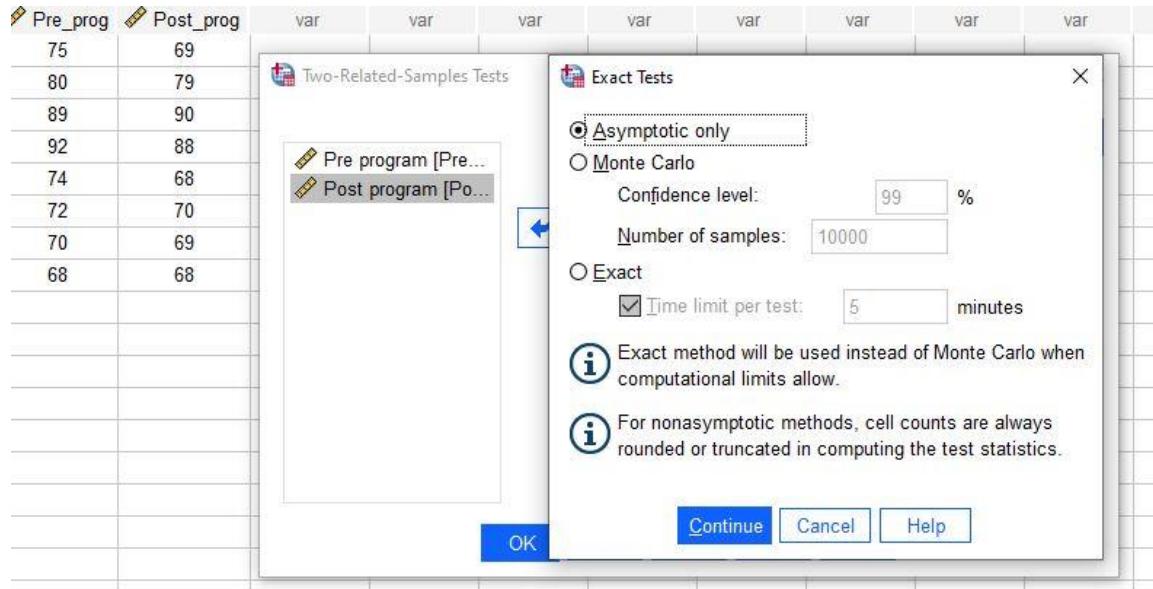
Figure 19.12: Exact test (Wilcoxon Signed Rank test).

Click on **Continue** and **OK** options to get the outputs:
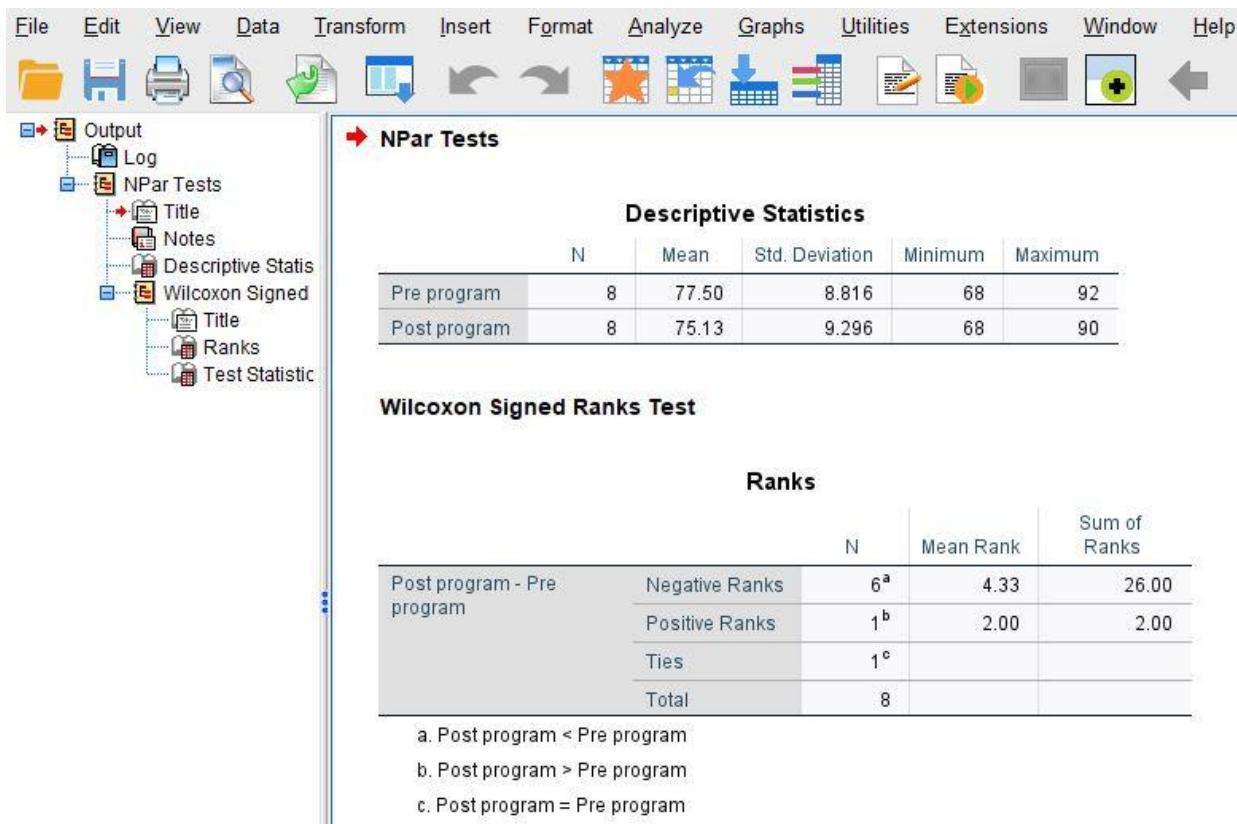
Figure 19.13: Descriptive Statistics and Ranks output (Wicoxon-Signed Rank test).

The second table (**Ranks**) provides information about the ranked scores (Figure 19.13). It tells us the number of **Negative Ranks** (these are people for whom the post-program weight was smaller than the pre-program weight) and the number of **Positive Ranks** (people for whom the post-program weight was greater than the pre-program weight). The table shows that for 6 out of the 8 students, their post-program weight was smaller than pre-test weight, indicating greater impact of gym training program compared to no program given. There was one **Ties** (i.e., participants who scored the same weight on both treatments). The table also shows the average number of negative and positive ranks and the sum of positive and negative ranks. Below the table are footnotes, which tell us to what the positive and negative ranks relate.

Figure 19.14: Test Statistics output (Wilcoxon-Signed Rank test).

Above table (Figure 19.14) tells us that the **Test Statistic** is based on the 'positive ranks', that the z-score is -2.514 (which is within rounding error of the value) and that this value is significant at $P = 0.04$. Therefore, because this value is based on the positive ranks (and because the test statistic is the smaller of the positive and negative ranks, most ranks must have been positive), we should conclude that when taking gym training program there was a significant decrease in post-program weight. If the **Test Statistic** had been based on the negative ranks then this would have told us that the results were in the opposite direction (i.e. post-program weight were larger as compared to pre-program). Therefore, we can conclude that for health program users there was a significant decrease in post-program weight; hence, gym training program was effective in reducing weight of the female students.

19.5: KRUSKAL–WALLIS T

In situations where the normality assumption is unjustified, the experimenter may wish to use an alternative procedure to the F test analysis of variance that does not depend on this assumption. Such a procedure has been developed by Kruskal and Wallis.

The Kruskal–Wallis test is a nonparametric alternative to the one-way ANOVA. It is used to compare three or more samples simultaneously and decide whether they belong to the same population or not. This test is used when the data obtained are measured at least on ordinal scale. The Kruskal–Wallis test can be used for parametric data if the assumption of normality does not hold or other assumptions required for one-way ANOVA violates.

Several nonparametric analogs to analysis of variance are available that use more information by considering the magnitude of each observation relative to the magnitude of every other observation. Perhaps the best known of these procedures is the Kruskal–Wallis one-way analysis of variance by ranks.

19.5.1: SPSS Procedure

University administration wanted to improve physical fitness for her students (Table 19.4). She offered three gym training programs with different intensities to choose from. In order to take decision, she conducted a study in which these programs were randomly allocated to the subjects from different degrees in the sample for 6 weeks. After the treatment was over, fitness score of the students was measured in three groups (1-10: 10 being the highest), which are shown in the following table. The normality assumption is violated here.

Table 19.4: Fitness score of students.

	Low	5	8	8	3	4	5	5	5	4	4
Program Intensity	**Medium**	6	6	6	2	5	6	4	4	7	6
	High	8	4	10	8	8	10	9	9	7	7

19.5.2: Q

Test variables included in this example are:
- *Test variable list = fitness score*
- *Grouping variable = training intensity* (3 *levels*: *low*; *medium*; *high*)

We need to test the following hypothesis:

$H_1 = At\ lesast\ one\ of\ the\ three\ program\ intensites\ differ\ in\ fitness\ score$

All the knowledge belongs to God.

To prepare data file, define Training as a nominal and Fitness as scale variables in the **Variable View** (Figure 19.15). For Training variable, define code 1 for Low Intensity, 2 for Medium Intensity, and 3 for High Intensity by clicking cell under the column heading **Values.**
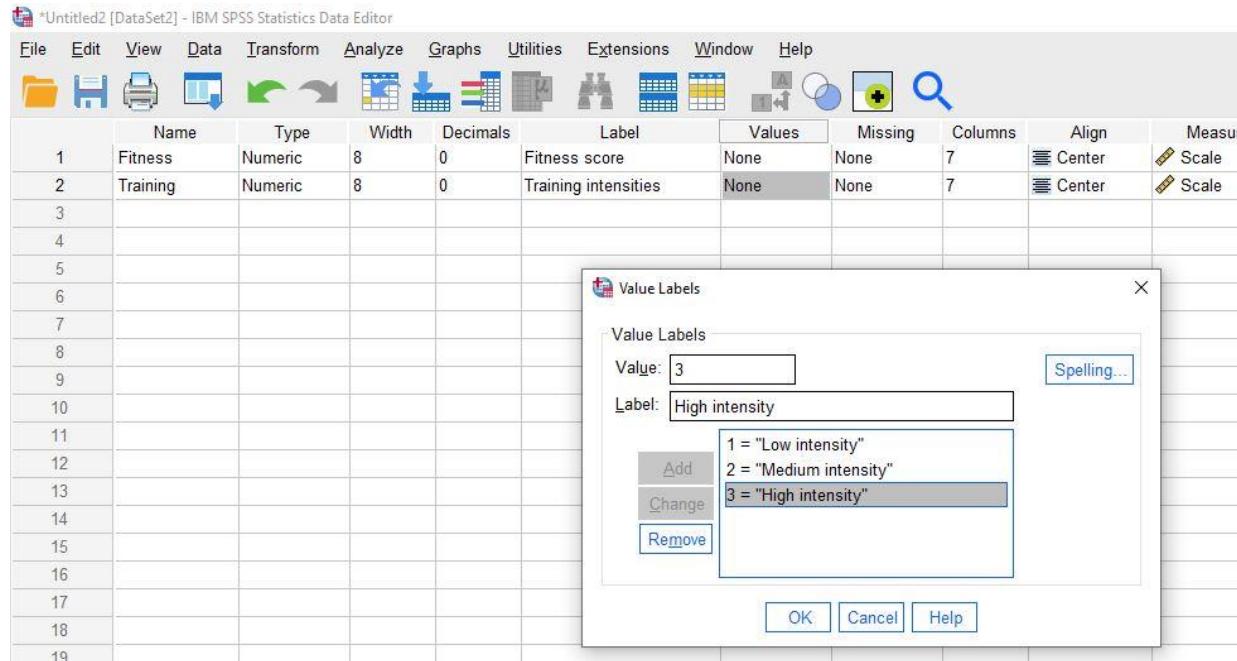
Figure 19.15: Defining variables (Kruskal-Wallis test).

In **Data View**, click on the following commands in sequence:
Analyze → Nonparametric Tests → Legacy Dialogs → K Independent Samples
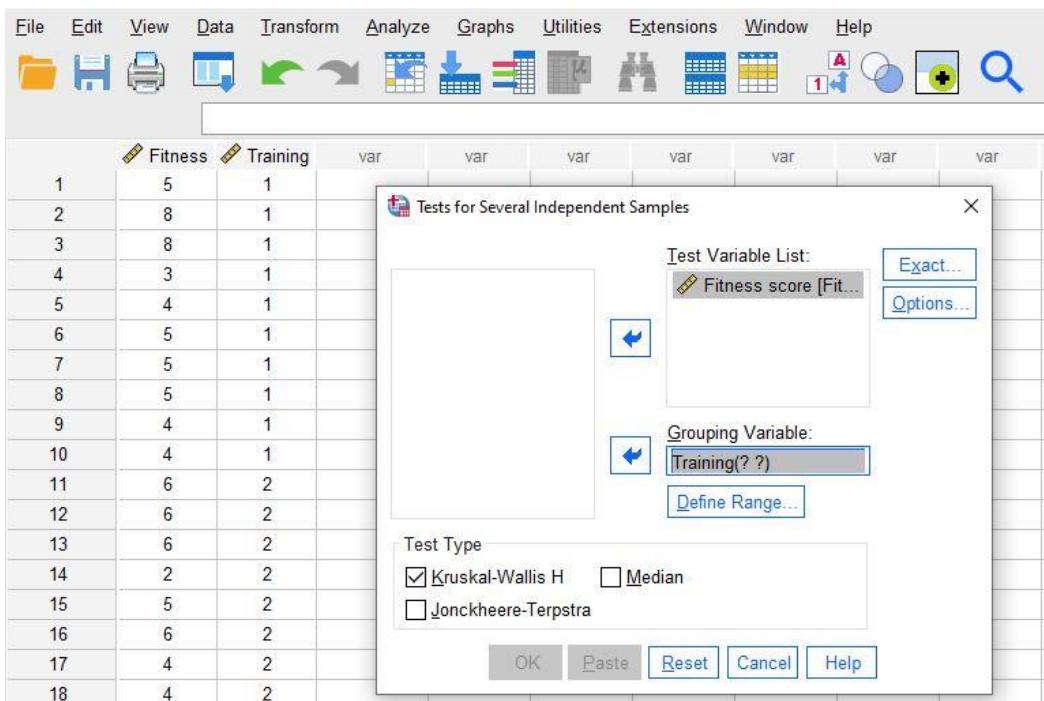
Figure 19.16: Variable input (Kruskal-Wallis test).

All the knowledge belongs to God.

After clicking **K Independent Samples** option, you will be directed to the screen as shown in Figure 19.16 for selecting variables and defining other options. Select Fitness and Training variables from the left panel and bring them into the **Test Variable List** and **Grouping Variable** sections in the right panel, respectively.

In **Define Range**, enter 1 in the box labeled **Minimum** and 3 in the **Maximum** as shown in Figure 19.17. Ensure that the **Kruskal–Wallis H** option is checked.
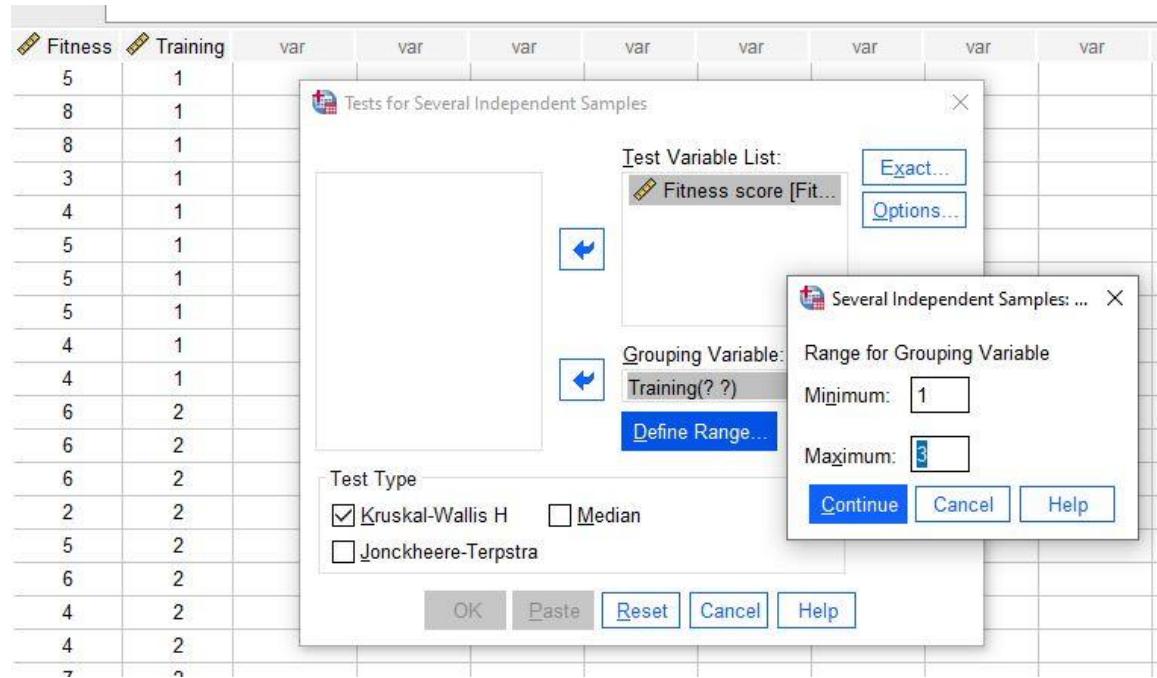
Figure 19.17: Define Range (Kruskal-Wallis test).

Click on **Options** command and check **Descriptive.** As mentioned previously, the **Exact** option can also be used (but not used here).
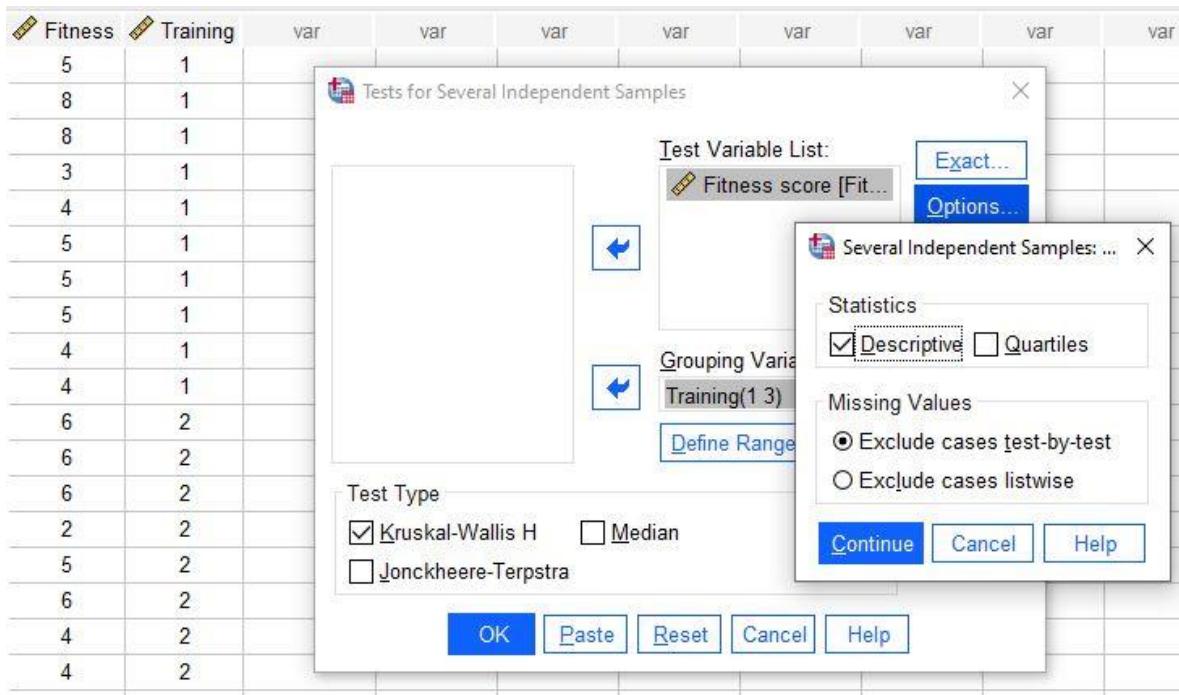
Figure 19.18: Options selection (Kruskal-Wallis test).

Click on **Continue** and **OK** options to get the output.
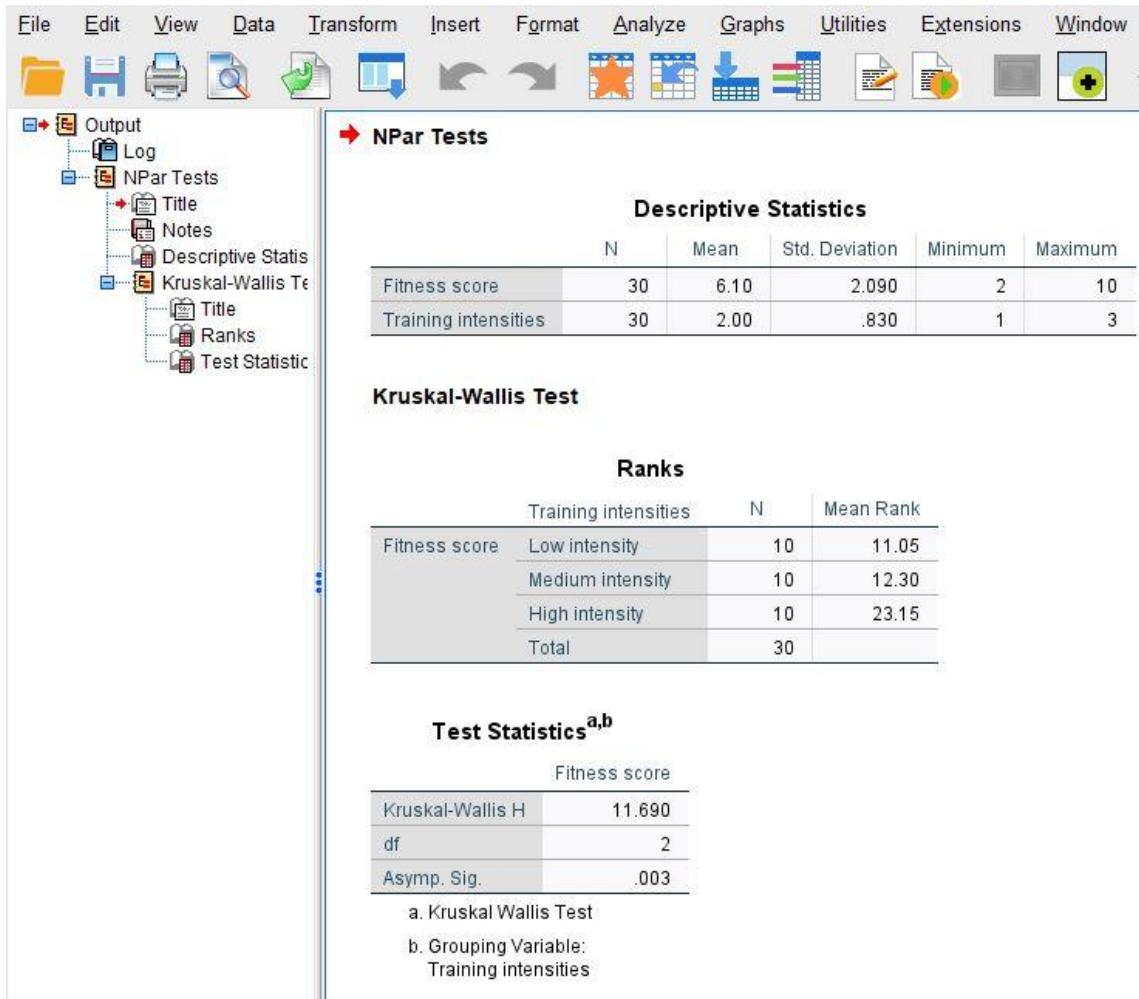
Figure 19.19: Descriptive Statistics, Ranks and Test Statistics output (Kruskal-Wallis test).

Since the asymptotic value ($Asym.Sig = 0.003$) from the Figure 19.19 is significant: ($p < 0.05$), it may be concluded that the Fitness Score differ significantly in the three treatment groups (Training Programs). And from the **Ranks** table it can be noted that the Fitness Score is highest for **High intensity** training program and lowest for the **Low intensity** program.

19.6: FRIEDMAN TEST

The Friedman test is a nonparametric test and can be used in place of one-way ANOVA with repeated measures if its assumptions violate. This test is used in a situation when repeated measures are obtained on the same set of subjects.

19.6.1: SPSS Procedure

A university wanted to improve her PhD students in research planning and thus offered Experimental Design as a subject during their course work. Eight students participated in the program. These students were tested in quiz for their knowledge before starting the course; after 2 weeks; in mid-term; in final exam, during the study period. The marks in quiz so obtained (1-10: 10 being the highest) is shown in Table 19.5. Since data violated the stringent assumptions of repeated measures ANOVA, apply suitable test.

All the knowledge belongs to God.

Table 19.5: Quiz marks during subject teaching

Time			
Day 1	**Week 4**	**Mid-term**	**Final**
1	4	6	8
1	4	6	8
0	5	6	7
0	3	4	6
2	3	4	8
2	8	8	10
3	2	3	6
1	2	4	6

Variable included in above problem is:

- $Test\ variables\ (within\ subject\ factor) = quiz\ (time)$

We need to test the following hypothesis:

$$H_1 = Change\ in\ time\ has\ an\ effect\ on\ quiz\ score.$$

When the data are collected using the same participants in each condition, the data being within-subjects factor, is entered using different columns. Define these variables in **Variable View** as following:

Figure 19.20: Defining variables (Friedman test).

In **Data View**, go to the following commands in sequence:

$$Analyze \rightarrow Nonparametric\ Tests \rightarrow K\ Related\ Samples$$

After clicking **K Related Samples** option, you will be taken to the screen as shown in Figure 19.21 for selecting variables and defining other options. Select Day_0, week_2, Mid and Final variables from the left panel and bring them into the **Test Variables** section in the right panel.

All the knowledge belongs to God.
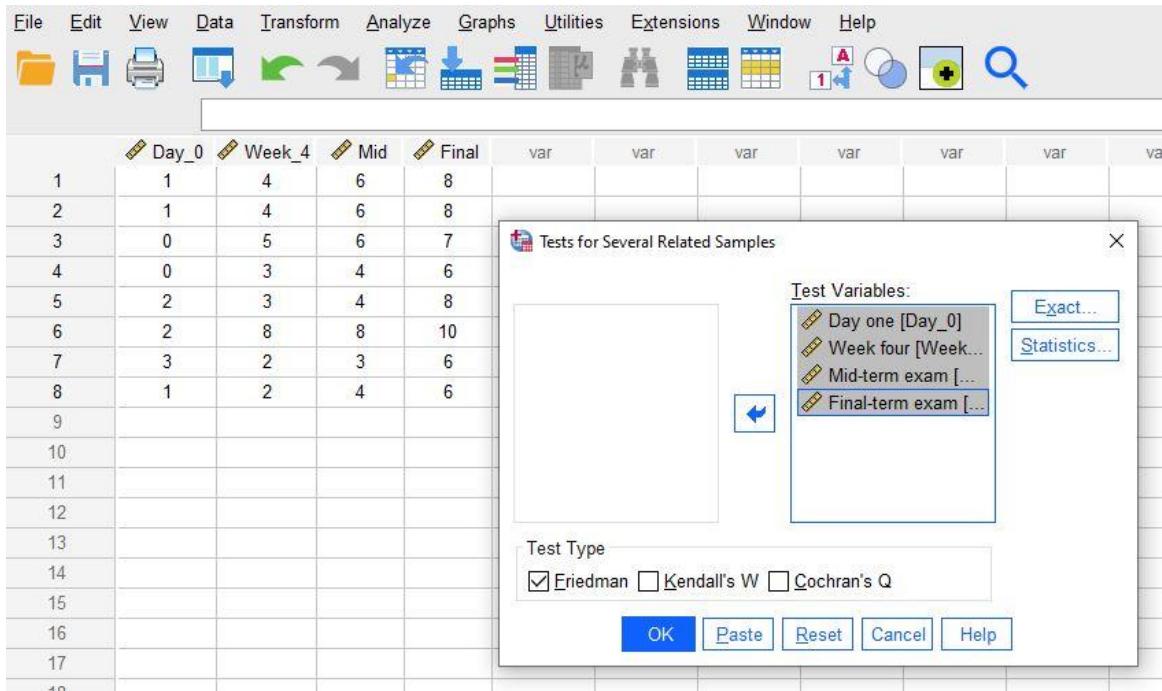
Figure 19.21: Variables input (Friedman test).

Click on **Statistics** command and check **Descriptive** (Figure 19.22). Ensure that the **Friedman** option is checked.
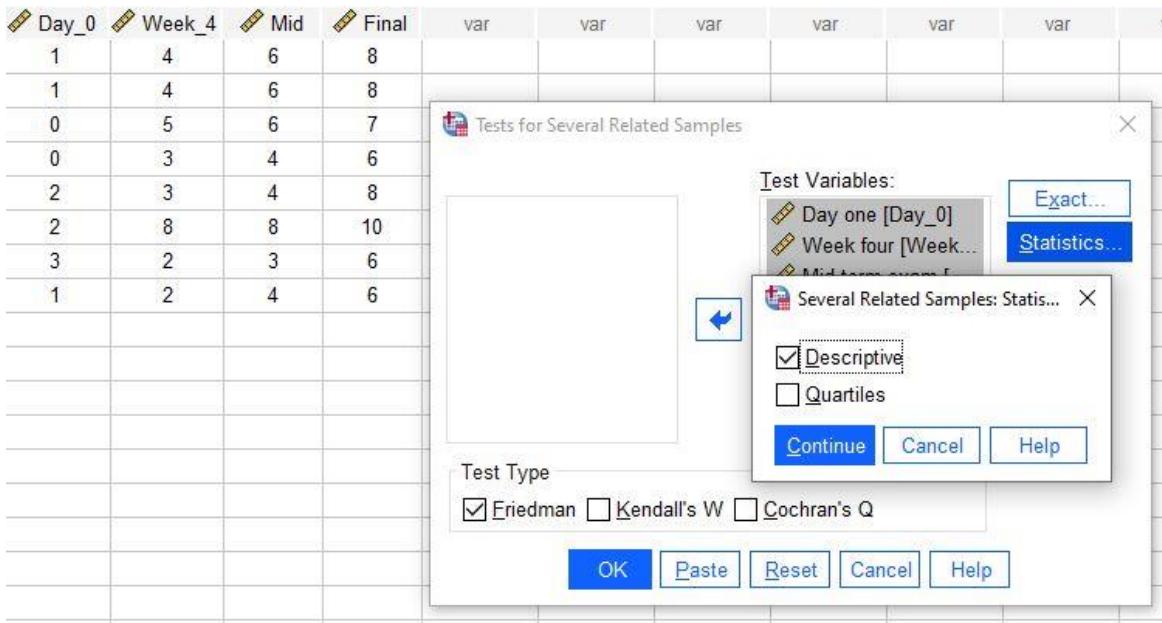
Figure 19.22: Statistics selection (Friedman test).

Click **OK** options to get the outputs:

All the knowledge belongs to God.
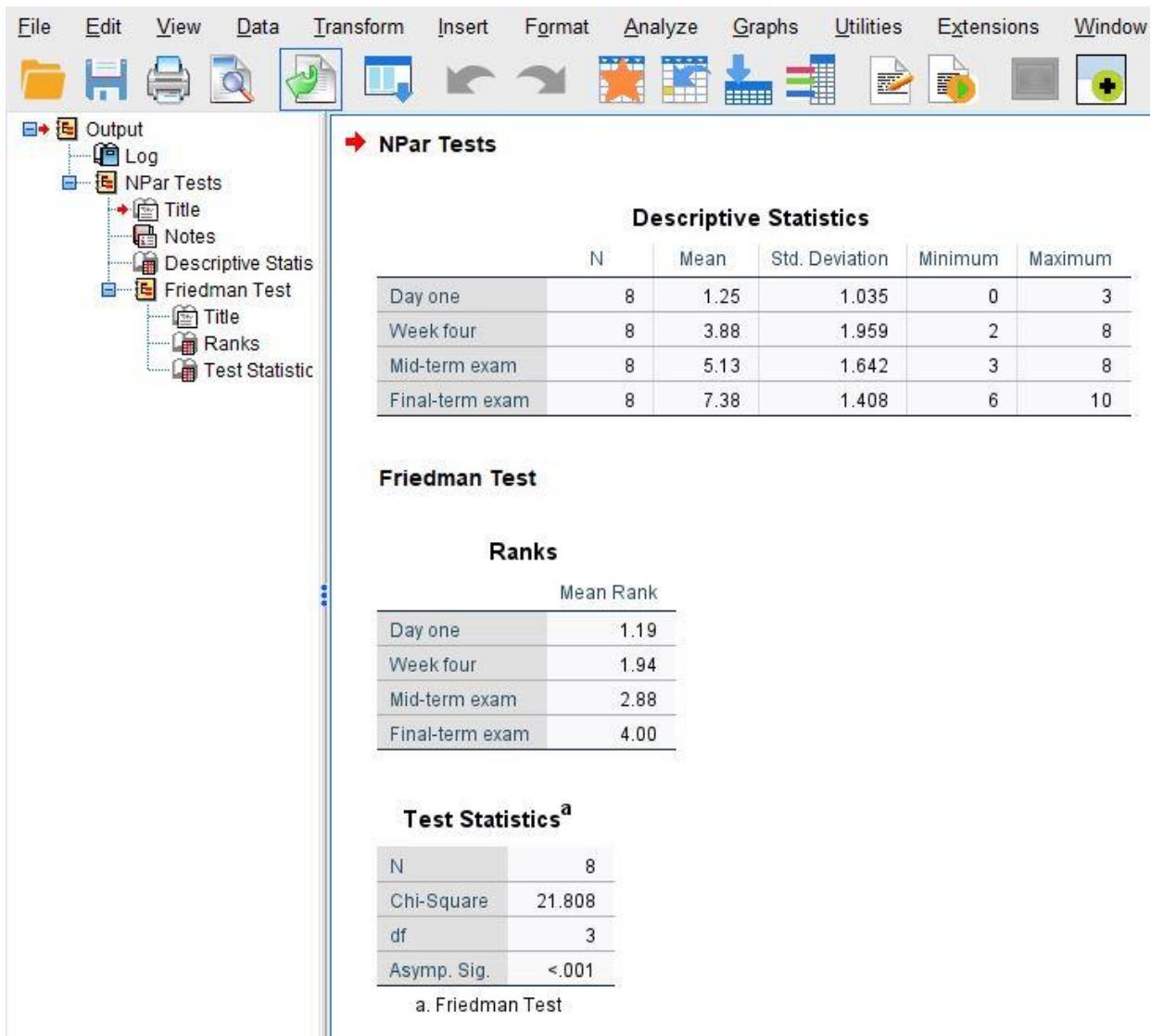
Figure 19.23: Descriptive Statistics, Ranks and Test Statistics output (Friendman test).

Descriptive Statistics (Figure 19.23) tells us that the highest mean is achieved in 'Final term exam' and lowest in 'Day one'. Similarly, **Ranks** table also indicate that **Mean Rank** of 'Final-term exam' is also highest among all.

Test Statistics table in Figure 19.23 is the point of focus here. **Asym.Sig** indicates that p < 0.001 is significant: (p < 0.05); hence, it may be concluded that the significant differences do exist in Quiz due to teaching of subject and that duration appears to increase marks of the students considerably.

CHAPTER 20: SOLVED EXERCISES

20.1: EXERCISE 1

20.1.1: Question 1

An experiment was carried out to know the effect of different levels of canola meal feeding (along with standard feed) on body weight in broiler chickens.

Table 20.1: Effect of canola meal on weight of broilers.

Canola meal (%)	Body weight after 30 days (grams)					
1	1802	1799	1834	1723	1811	
3	1912	1934	1985	1954	1987	1977
5	2143	2143	2143	2193	2188	
7	2423	2453	2477	2412	2423	
9	2013	2076	2098	2043		

Recognize the design along with its components (hypothesis; variables; model).

20.1.1.1 Solution

Design will be CRD and one-way ANOVA will be used for data analysis. Variables included in this example are:

- $Factor = canola\ meal\ feeding$
- $Levels\ of\ factor = five$
- $Response = body\ weight$
- $EUs = broiler\ chickens$

The hypothesis needs to be tested is:
$$H_1 = Change\ in\ canola\ feeding\ has\ an\ effect\ on\ broiler\ weight$$

Model will be:
$$Broiler\ weight = canola\ meal + Error$$

20.1.2: Question 2

Let us suppose we are to conduct an experiment taking 6 treatments (A, B, C, D, E, F) in 3 replications (R1, R2, R3). Six treatments are to be allocated randomly among the six EUs of each block separately:

Table 20.2: Six treatments under 3 replications.

R1: Replication 1	R2: Replication 2	R3: Replication 3
E	B	C
C	C	E
B	E	F
D	D	A
A	F	D
F	A	B

Recognize the design along with its components (hypothesis; variables; model).

20.1.2.1 Solution

RCB design will be used. Variables included in this problem are:

- $Factor = six\ treatments$

All the knowledge belongs to God.

- $Constraint = replications\ (3\ levels)$
- $Response = X$
- $EUs = Y$

Hypotheses to be tested are:
$$H_1 = Change\ in\ treatment\ has\ an\ effect$$

$$H_1' = Change\ in\ replication\ has\ an\ effect$$

Model will be:
$$X = (treatment + replication) + Error$$

20.1.3: Question 3
Suppose we are to test 5 new chicken varieties (viz., C1, C2, C3, C4, C5) along with a traditional check (Ross) for their survivability under a particular local situation in Islamabad. Suppose we are provided with material such that there could be 5, 4, 4, 3, 3, and 5 replications for C1, C2, C3, C4, C5, and Ross, respectively. The experimental house is required to be divided into 24 (5 + 4 + 4 + 3 + 3 + 5) experimental units of equal size and preferably of equal shape with all materials supposed to be homogenous. Recognize the design with its components.

20.1.3.1 Solution
CR is design and one-way ANOVA will be used in SPSS. Variables are:
- $Factor = chicken\ varities$
- $Levels\ of\ factor =\ six\ variteis\ (5\ new\ varieties + 1\ traditional)$
- $Response = survivability$
- $EUs = chickens$

Hypothesis:
$$H_1 = Change\ in\ chicken\ variety\ has\ an\ effect\ on\ survivability$$

Model:
$$Survivability = chciken\ varieties + Error$$

20.1.4: Question 4
A random sample of 20 university students was tested for their performance in quiz. Their scores (out of 50) were: 23, 26, 25, 26, 20, 29, 27, 27, 25, 26, 28, 25, 29, 29, 29, 29, 23, 28, 35 and 31. Can it be concluded that all the university students have mean scores equal to 30?

20.1.4.1 Solution
One sample T test will be used. Variables included are:
- $Test\ variable\ =\ quiz\ performance$
- $Test\ value =\ 30$
- $Response = test\ scores$
- $EUs = university\ students$

Hypothesis:
$$H_1 = The\ mean\ quiz\ score\ is\ different\ than\ 30$$

All the knowledge belongs to God.

20.1.5: Question 5

A scientist was interested in investigating the effect of different types of feeds used by farmers on broiler body weight. Three groups of broilers were randomly made with equal initial weight, and they were offered three different types of feeds (SB, SP, and ISB) in equal amount during three different rearing. Before, during each week and after the trial, their body weights were measured and the gain in weight in kgs was recorded.

Formulate design.

20.1.5.1 Solution

Repeated measures ANOVA will be used for design formulation. Variables are:

- $Factor = feed$
- $Levels\ of\ factor = 3\ (SB; SP; ISB)$
- $Response = weight\ gain$
- $EUs = broilers$

Hypothesis:
$$H_1 = At\ least\ one\ of\ the\ feed\ has\ diferent\ response\ than\ others$$

20.1.6: Question 6

In fish breeding improvement program, fishes are subjected to five different light intensities. The following table gives the fish weight (in kg) at the harvest after 90 days. Test whether the different light intensities differ significantly with respect to fish weight:

Table 20.3: Effect of different light intensities on fish weight.

T1	T2	T3	T4	T5
2.2	1.5	1.8	2.1	2.2
2.3	1.5	1.9	2.4	2.3
2.2	1.6	1.8	2.3	2.4
2.3	1.6	1.9	2.4	2.4
2.2	1.6	1.8	2.3	2.3

20.1.6.1 Solution

Design will be CRD, and one-way ANOVA can be used in SPSS for analysis. Variables included are:

- $Factor = light\ intensities$
- $Levels\ of\ factor = 5\ (T1; T2; T3; T4; T5)$
- $Response = fish\ weight$
- $EUs = fish$

Hypothesis:
$$H_1 = Change\ in\ light\ intensities\ has\ an\ effect\ on\ fish\ weight$$

Model:
$$Fish\ weight = light\ intensities + Error$$

All the knowledge belongs to God.

20.1.7: Question 7

A field experiment was conducted to study the effect of 8 different feeds on the weight gain (grams) of hybrid broiler. The following is the layout and data pertaining to weight gain at 7 days after feeding:

Table 20.4: Weight gain of broilers.

Rep-1	Rep-2	Rep-3
T4 (96)	T1 (180)	T6 (197)
T3 (145)	T3 (140)	T8 (339)
T5 (196)	T8 (335)	T3 (147)
T7 (79)	T4 (95)	T5 (196)
T6 (197)	T6 (193)	T4 (98)
T2 (169)	T2 (174)	T7 (87)
T8 (338)	T7 (86)	T2 (174)
T1 (182)	T5 (187)	T1 (184)

Recognize the design along with its components (hypothesis; variables; model).

20.1.7.1 Solution

Design name will be RCBD (as blocking has been used in the form of 3 replications) and two-way ANOVA will be used for analysis. Variables included are:

- $Factor\ 1 = feeds\ (8\ levels)$
- $Blocking = replications\ (3)$
- $Response = weight\ gain$
- $EUs = broilers$

Hypotheses will be:

$$H_1 = Change\ in\ feed\ has\ an\ effect\ on\ weight\ gain$$

$$H_1' = Change\ in\ replication\ has\ an\ effect\ on\ weight\ gain$$

Model:

$$Weight\ gain = (feed + replication) + Error$$

20.1.8: Question 8

Table 20.5 shows the results of an experiment on deer. They were divided into 30 groups of 100 deer each. The 30 groups were divided at random into 10 groups of 3 pens each, with each group receiving a different feed treatment. The ten treatments are the combinations of amount of feed used with 2 levels, and age of the deer with 5 levels. The response measured is the weight gain (kgs).

All the knowledge belongs to God.

Table 20.5: Deer body weight under different treatments.

Feed	Age of deer (weeks)				
	10	12	14	16	18
Anie feeds	11	7	9	13	20
	9	16	19	35	37
	6	17	35	28	45
Lahore feeds	8	1	5	1	11
	3	7	9	10	15
	3	3	9	9	25

Recognize the design along with its components (hypothesis; variables; model).

20.1.8.1 Solution

It is an RCBD under factorial treatment and factorial ANOVA (two-way ANOVA) will be used for analysis. Variables are:

- $Factor\ 1 = feed\ (2\ levels)$
- $Factor\ 2 = deer\ age\ (5\ levels)$
- $Response = deer\ weight\ gain$
- $EUs = deer$

Hypotheses need to be tested are:

$$H_1 = Change\ in\ feed\ has\ an\ effect\ on\ weight\ gain$$

$$H_1' = Change\ in\ deer\ age\ has\ an\ effect\ on\ weight\ gain$$

$$H_1'' = There\ exist\ an\ interaction\ between\ deer\ age\ and\ feed\ type$$

Model:

$$Weight\ gain = (age + feed) + (age \times feed) + Error$$

20.2: EXERCISE 2
20.2.1: Question 1

Table 20.6: Starter egg weight (g) from a layer with 9 layer-feed treatments.

	R-1	R-2	R-3
T1	28	27	31
T2	29	29	29
T3	30	29	30
T4	30	31	29
T5	28	29	30
T6	28	26	27
T7	32	29	30
T8	32	31	29
T9	27	29	29

All the knowledge belongs to God.

Recognize the design along with components of design formulations.

20.2.1.1 Solution

RCB design will be used due to replication as blocking. Two-way ANOVA will be used for analysis. Variables included are:

- $Factor\ 1 = feed\ treatment\ (9\ levels)$
- $Blocking = replication\ (3\ levels)$
- $Response = egg\ weight$
- $EUs = layer\ chickens$

Hypotheses:
$$H_1 = Change\ in\ feed\ has\ an\ effect\ on\ egg\ weight$$

$$H_1' = Change\ in\ replication\ has\ an\ effect\ on\ egg\ weight$$

Model:
$$Egg\ weight = (feed + replication) + Error$$

20.2.2: Question 2

Ali studied the effect of acid rain on trees in Margla (Islamabad). Clean precipitation has a pH in the 5.0 to 5.5 range, but observed precipitation pH in Islamabad is often in the 3.0 to 4.0 range. Is this acid rain harming trees, and if so, does the amount of harm depend on the pH of the rain?

One of their experiments used 240 six-week-old seedlings. These seedlings were divided into five groups of 48 at random, and the seedlings within each group received an acid mist treatment 6 hours a week for 17 weeks. The five treatments differed by mist pH: 4.7, 4.0, 3.3, 3.0, and 2.3; otherwise, the seedlings were treated identically. After the 17 weeks, the seedlings were weighed, and total plant (dry) weight was taken as response.

Recognize the design along with its components and variables.

20.2.2.1 Solution

Design name will be CRD. Variables are:

- $Factor\ 1 = acid\ rain\ pH$
- $Levels\ of\ factor = 5\ treatment\ pH\ (4.7;\ 4.0;\ 3.3;\ 3.0;\ 2.3)$
- $Response = plant\ weight$
- $EUs = birch\ seedlings$

Hypothesis:
$$H_1 = Change\ in\ pH\ has\ an\ effect\ on\ plant\ weight$$

Model:
$$Plant\ weight = pH + Error$$

Layout:

All the knowledge belongs to God.

Table 20.7: Plant yield under 5 types of acid rain pH.

Acid rain pH				
4.7	**4.0**	**3.3**	**3.0**	**2.3**
Plant height	Plant height	Plant height	Plant height	Plant height
Plant height	Plant height	Plant height	Plant height	Plant height
Plant height	Plant height	Plant height	Plant height	Plant height

20.2.3: Question 3

To know the effect of different levels of pond water (PW) and sunlight on fish yield, an experiment with 2 levels of pond water (50 inch and 100 inch) and 2 levels of sunlight (500 lux and 10,000 lux) was laid with 5 replications. The Table 20.8 gives the fish yield in kg.

Analyze the data to find out (i) the best level of water, (ii) the best amount of sunlight, and (iii) the combination of water and sunlight for the best yield of fish:

Table 20.8: Fish yield in 5 replications.

Fish yield (kg)									
	R1		**R2**		**R3**		**R4**		**R5**
PW	500 lux	10,000 lux	500 lux	10,000 lux	500 lux	10,000 lux	500 lux	10,000 lux	500 lux	10,000 lux
50 inch	1.81	1.97	1.78	1.95	1.83	1.99	1.87	1.96	1.77	1.93
100 inch	2.17	2.42	2.19	2.37	2.22	2.4	2.07	2.47	2.15	2.41

20.2.3.1 Solution

Design will be RCBD under factorial and factorial ANOVA (two-way) will be used for analysis. Variables are:

- $Factor\ 1 = nitrogen\ (2\ levels)$
- $Factor\ 2 = sulfur\ (2\ levels)$
- $Response = garlic\ yield$
- $EUs = garlic$

Hypotheses need to be tested are:

$$H_1 = Change\ in\ sulfur\ dose\ has\ an\ effect\ on\ garlic\ yield$$

$$H_1' = Change\ in\ nitrogen\ dose\ has\ an\ effect\ on\ garlic\ yield$$

$$H_1'' = There\ exist\ an\ interaction\ between\ nitrogen\ and\ sulfur\ doses$$

Model:

$$garlic\ yield = (nitrogen + sulfur) + (nitrogen \times sulfur) + Error$$

20.2.4: Question 4

In a trail carried out with five strains of chicken of a particular breed, each bird received five different feeds in five successive periods. Egg weights (in grams) were calculated as follows:

All the knowledge belongs to God.

Table 20.9: Egg weights (g) of chickens.

Chicken breeds	Feeds				
	1	2	3	4	5
1	60 B	67 D	63 E	54 A	62 C
2	54 C	64 A	63 B	64 E	71 D
3	66 D	63 C	60 A	65 B	63 E
4	65 E	66 B	64 C	70 D	52 A
5	67 A	65 E	73 D	53 C	60 B

Recognize the design and analyze data.

20.2.4.1 Solution

It is symmetrical LS design and three-way ANOVA will be used for analysis. Variables included are:

- $Factor = feeds$ (5 *levels*)
- $Blocking\ 1 = breeds$ (5 *levels*)
- $Blocking\ 2 = periods$ (5 *levels*)
- $Response = egg\ weight$
- $EUs = chickens$

Hypotheses:

$$H_1 = Change\ in\ feed\ has\ an\ effect\ on\ egg\ weight$$

$$H_1' = Change\ in\ chicken\ breed\ has\ an\ effect\ on\ egg\ weight$$

$$H_1'' = Change\ in\ periods\ has\ an\ effect\ on\ egg\ weight$$

$$H_1''' = There\ exist\ an\ interaction\ between\ feed, breed\ and\ periods$$

We will use the following keys to avoid the lengthy model:

- $Chicken\ breeds = C$
- $Feed = F$
- $Periods = P$
- $Egg\ weight = EW$

So, the model will be as following:
$$EW = (F + P + C) + (F \times P \times C) + (F \times P) + (F \times C) + (C \times P) + Error$$

All the knowledge belongs to God.

20.2.5: Question 5

Table 20.10: The egg moisture loss data (%) from a trial with 7 chicken varieties.

	Egg moisture loss (%)			
Varieties	Block 1	Block 2	Block 3	Block 4
P1	6	8	7	7
P2	9	10	10	10
P3	8	8	8	8
P4	9	9	10	10
P5	9	9	10	10
P6	8	8	7	8
P7	8	8	9	9

Recognize the design along with components of design formulations.

20.2.5.1 Solution

It is a RCB design due to blocking as factor and two-way ANOVA will be used for analysis. Variables included are:

- $Factor = chicken\ varieties\ (7\ levels)$
- $Blocking = blocks\ (4\ levels)$
- $Response = moisture\ loss\ \%$
- $EUs = chickens$

Hypotheses are:

$H_1 = Change\ in\ chicken\ variety\ has\ an\ effect\ on\ egg\ moisture\ loss$

$H_1' = Change\ in\ block\ has\ an\ effect\ on\ egg\ moisture\ loss$

Model:

$Egg\ moisture\ loss = (chciken\ variety + blocks) + Error$

20.2.6: Question 6

A poultry scientist developed an experimental process to purify drinking water for birds, to be useful, must not change the acidity of the treated water (ideally it would maintain a neutral pH of 7.0). To assess the process, mean of a sample of pH-values is compared to the hypothetical mean pH of 7.0. The experimental process applied many times and observations were recorded which are assumed to be independent and normally distributed.

Table 20.11: pH values of water after treatment.

5.9	7.3	6.8	6.5	6.7	6.6	6.9
7.5	6.6	6.9	6.9	6.3	7.2	6.3
6.5	6.6	6.9	6.9	7.2	6.8	6.8

Analyze the data and check the pH of water.

All the knowledge belongs to God.

20.2.6.1 Solution

One sample T test will be used. Variables included in this example are:

- $Test\ variable = water\ pH$
- $Test\ value = 7.0\ pH$

Hypothesis:
$$H_1 = The\ pH\ value\ of\ water\ is\ not\ 7.0$$

20.2.7: Question 7

For $t = 5$ and $p = 3$ the following design is a balanced design with 10 subjects in which each treatment is preceded and followed by every other treatment exactly once:

Table 20.12: Treatments applied on 10 subjects in different time periods.

Period	Subject									
	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	3	4	5	1	2
2	5	1	2	3	4	5	1	2	3	4
3	3	4	5	1	2	1	2	3	4	5

Recognize the design.

20.2.7.1 Solution

It is Crossover design with LSD as basic one. Factorial treatment will not be involved as both time and subject have been taken as blocking here. Variables included are:

- $Factor = treatment\ (5\ levels)$
- $Blocking\ 1 = time\ period\ (3\ levels)$
- $Blocking\ 2 = subjects\ (10\ levels)$
- $Response = X$
- $EUs = Y$

Hypotheses will be:
$$H_1 = Change\ in\ treatment\ has\ an\ effect$$

$$H_1' = Change\ in\ subject\ has\ an\ effect$$

$$H_1'' = Change\ in\ time\ period\ has\ an\ effect$$

Model:
$$X = (treatment + time\ period + subjects) + Error$$

20.2.8: Question 8

A field experiment was conducted to identify the best spacing and lighting. Three different spacing were randomly allocated to three poultry houses in each replication separately, and in each house, three lights were allocated randomly among the four pens in each house. Broiler yields (grams) were recorded from the individual pen and given below. Analyze the data and draw your conclusion:

All the knowledge belongs to God.

Table 20.13: Broiler yield due to spacing and lighting.

Spacing	Spacing-1				Spacing-2				Spacing-3			
House	P1	P2	P3	P4	P1	P2	P3	P4	P1	P2	P3	P4
1	17.4	11.8	9	7.7	12	10	4	3	10	4	2.9	2.2
2	17.45	12	8	8	11.9	10.1	4	3	9	4.1	3	3
3	17.4	11.9	8	7.7	11.9	10	4	3.3	9.7	4.1	2.9	2

20.2.8.1 Solution

It is a split-plot design, and three-way ANOVA will be used for analysis. Variables included are:

- $Main\ plot\ factor = spacing\ (3\ levels)$
- $Sub\ plot\ factor = penning\ (4\ levels)$
- $Blocking = houses\ (3\ levels)$
- $Response = broiler\ yield$
- $EUs = chickens$

Hypotheses:

$$H_1 = Change\ in\ spacing\ has\ an\ effect\ on\ yield$$

$$H_1' = Change\ in\ penning\ has\ an\ effect\ on\ yield$$

$$H_1'' = Change\ in\ house\ has\ an\ effect\ on\ yield$$

$$H_1'' = There\ exist\ an\ interaction\ between\ house, penning\ and\ spacing$$

To avoid the lengthy model, we will use following key:

- $Spacing = S$
- $Penning = P$
- $Housing = H$
- $Broiler\ yield = BY$

So, the model will look like:
$$BY = (S + P + H) + (S \times P \times H) + (S \times P) + (S \times H) + (P \times H) + Error$$

20.3: EXERCISE 3

20.3.1: Question 1

Consider the blood concentration of amoxicillin after it has been administered to deer. The concentration will typically start at zero, increase to some maximum level as the drug gets into the bloodstream, and then decrease back to zero as the drug is metabolized or excreted. These time-concentration curves may differ if the drug is delivered in a different form, say a syrup versus an injection. Bioequivalence studies seek to determine if different drug delivery systems have similar biological effects. One variable to compare is the area under the time-concentration curve. This area is proportional to the average concentration of the drug.

We wish to compare three methods for delivering a drug: a solution, a tablet, and an injection. Our response will be the area under the time-concentration curve. We anticipate large subject to subject differences, so we block on subject. There are three subjects, and each subject will be given the drug three times, once with each of the three methods. Because the body may adapt to the drug in some way, each drug will be used once in the first period, once in the second

All the knowledge belongs to God.

period, and once in the third period. Following table gives the assignment of treatments and the responses.

Table 20.14: Time-concentration curve due to different drug delivery methods.

		Subject		
		1	2	3
	1	A 1799	C 2075	B 1396
Period	2	C 1846	B 1156	A 868
	3	B 2147	A 1777	C 2291

20.3.2: Question 2

During a weight gain study, each of nine desi chickens was given probiotics for 2 weeks and then a placebo for another 2 weeks, or the placebo for the first 2 weeks and then probiotics for the second 2 weeks. As part of the study, the subjects were weighed to know the response at the end of each 2-week period. The weight gain data are shown in the following table:

Table 20.15: The weight gain of broilers.

	Weight gain (g)	
Chickens	Probiotics	Placebo
1	79	78
2	48	54
3	52	142
4	15	25
5	61	101
6	107	99
7	77	94
8	54	107
9	5	64

20.3.3: Question 3

Sales data of Centaurus Mall, Islamabad (% increase in sales) for four displays is shown in following table. Each display was installed in 5 different floors for 1 week:

Table 20.16: Effect of stores on sales %.

Displays			
D1	D2	D3	D4
4.2	8.4	3	4.9
2.7	4.5	3.8	2.8
3.1	4.9	2	6.1
4.6	7.3	2.1	4.2
1.2	5.7	3.2	3.7

Analyze data as per design.

All the knowledge belongs to God.

20.3.4: Question 4

Recognize the following design and analyze data.

Table 20.17: Different procedures of preparing fish feed.

		Column						
		1	2	3	4	5	6	7	8
Rows	1	17 D3	17 H4	17 C5	17 B6	16 E8	18 A1	16 G2	16 F7
	2	16 F6	16 E5	16 G4	19 A3	18 H2	18 B7	19 C8	17 D1
	3	17 B5	17 C6	19 H3	17 D4	16 G1	18 F8	19 E7	18 A2
	4	18 A4	16 G3	16 E6	16 F5	18 C7	18 D2	18 H1	17 B8
	5	17 C1	17 B2	17 D8	19 H7	20 A5	19 E3	16 F4	16 G6
	6	17 E2	16 F1	17 A7	16 G8	17 D6	18 C4	17 B3	20 H5
	7	16 G7	17 A8	16 F2	17 E1	18 B4	19 H6	17 D5	19 C3
	8	19 H8	17 D7	17 B1	19 C2	17 F3	16 G5	17 A6	18 E4

20.3.5: Question 5

Consider an experiment on the survival of *E coli*, in which three factors were examined. Three levels of acetic acid, six levels of litter depth and three yogurt levels were combined to give different combinations. The data is presented in the following table:

All the knowledge belongs to God.

Table 20.18: Survival rate of salmonella experiment.

Acetic acid (ppm)	Yogurt (g/kg)	Litter depth (inch)					
		1	2	3	4	5	6
0	2	4.2	4.5	5	6.1	6.2	8.3
	4	4.3	4.3	5.3	5.9	6.7	8.3
	6	4.3	4.8	5	5.8	6.6	8.1
20	2	4.1	4.1	4.2	5.7	6.5	7.5
	4	4.3	4.4	4.9	5.2	6.1	7.7
	6	4.1	4.2	4.2	5	6.5	7.6
40	2	4.1	4.3	4.7	5.4	6.4	7.1
	4	4.1	4.2	4.4	5.1	6.1	6.9
	6	3.9	4.2	4.4	5.2	6.3	7.1

20.3.6: Question 6

Chickens can be infested with Northern Mite, and the farmer wishes to test 3 treatments: water (a control), kerosene oil and pyrethrin. Five infested chickens are removed to a testing area. Left wing feathers are randomly chosen on each chicken, and two patches are marked on each wing; the number of mites in these patches is noted. Treatment is applied on these patches. After 3 days, the patches are counted again, and the response is the change in the number of mites (in following table).

Recognize the design.

Table 20.19: Effect of insecticides on mites on infested chicken.

	Chickens				
	1	2	3	4	5
Water	-9	18	10	9	-6
	-6	5	9	0	13
Kerosene oil	-4	29	4	-2	11
	7	10	-1	6	-1
Pyrethrins	4	29	14	14	7
	11	36	16	18	15

20.3.7: Question 7

An experiment was conducted to assess the best light intensity (lux) and space (square feet) in broiler breeder to get maximum egg production. The following data gives the yield in percentage. Analyze the data to estimate the best light intensity and space along with the best combination of these factors to provide maximum yield.

All the knowledge belongs to God.

Table 20.20: Egg production % in broiler breeders.

	Light intensity	
Space	50 lux	80 lux
2.0 sq.ft	85	88
	84	90
	85	91
2.5 sq.ft	85	94
	86	94
	86	93

20.3.8: Question 8

A study was conducted on students to compare the flexibility (1-50; 50 being highest) of male, female, and transgenders. The data so obtained is shown in following table. Apply test and discuss your findings at 5% level.

Table 20.21: Flexibility score.

Male	10	14	11	15	13	13	10	9	10	15
Female	18	17	15	16	16	18	17	17	13	17
Transgender	12	15	11	13	12	17	13	13	14	14

20.3.9: Solution

- Question 1: Crossover LSD.
- Question 2: Crossover design.
- Question 3: CRD.
- Question 4: GLSD.
- Question 5: Factorial ANOVA (three-way).
- Question 6: RCBD or two-way ANOVA.
- Question 7: Factorial design.
- Question 8: CRD or one-way ANOVA.

20.4: EXERCISE 4

20.4.1: Question 1

The objective of this experiment is to define effects of the following factors on the corrosion rate of pipes used in water system: acetic acid; organic debris (A; B; C; D; E) ; size of pipe; water pH (1; 2; 3; 4; 5). Five pipes were exposed to factors until a change in its color from colorless to light brown occurred. The corrosion intensity measure is the loss in weight of pipes (micro grams) after rinsing and drying them. The results of the experiment are shown in Table 20.22.

Recognize the design and analyze data.

All the knowledge belongs to God.

Table 20.22: Corrosion intensity measure of water pipes.

Size of pipe (inch)	Acetic acid (ppm)				
	10	**50**	**100**	**150**	**200**
1	65 A1	82 B3	108 C5	101 D2	126 E4
2	84 B2	109 C4	73 D1	97 E3	83 A5
3	10 C3	129 D5	89 E2	89 A4	52 B1
4	11 D4	72 E1	76 A3	117 B5	84 C2
5	97 E5	59 A2	94 B4	78 C1	106 D3

20.4.2: Question 2

A remotivation team conducted an experiment to compare five methods for remotivating students. Students were grouped according to level of initial motivation. Students in each group were randomly assigned to the five methods. At the end of the experimental period the students were evaluated by a team composed of a psychiatrist, a motivational speaker, a teacher, and a social worker, none of whom was aware of the method to which patients had been assigned. The team assigned each student a composite score as a measure of his or her level of motivation. The results are in following table.

Analyze the data.

Table 20.23: Composite score as a measure of level of motivation.

Level of Initial Motivation	Re-motivation Method				
	A	**B**	**C**	**D**	**E**
Nil	58	68	60	68	64
Very low	62	70	65	80	69
Low	67	78	68	81	70
Average	70	81	70	89	74

20.4.3: Question 3

A psychologist was interested to investigate the effect of exam on the stress level of DVM students. He conducted a study on randomly selected five male students. All subjects were tested for their stress level under each of the three exam levels (tough, medium and easy). The data so obtained is shown in following table.

Recognize the design and analyze data.

Table 20.24: Data on stress under all exams.

Cognitive theory		
Tough	**Medium**	**Easy**
35	36	38
32	28	41
29	31	40
28	34	42
30	33	39

20.4.4: Question 4

A farmer conducted a field trial to compare the relative effects of 5 particular feeds on the yield of broiler chicken. Thirty homogeneous experimental units are available and 6 were assigned at random to each feed treatment. At harvest time, 3 samples were taken

All the knowledge belongs to God.

(at random) from each experimental unit and the yield was obtained for each of the chickens. Recognize the design.

20.4.5: Question 5

Table 20.25: Weight gain affected by 3 types of feed and 3 temperatures.

Feed	Temperature (°F)		
	55	70	80
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	159	106	58
	188	122	70
	126	115	45
3	138	174	96
	168	150	82
	110	120	104
	160	139	60

20.4.6: Question 6

University students were questioned about the number of hours they sleep each day. We want to test the hypothesis that the DVM students sleep less than the Zoology students. To test this claim following data is collected.

Table 20.26: Sleep response in students.

DVM	6.2	6.8	7.2	7.9	8.1	5.1	5.5	5.9	6.1
Zoology	8	8.1	9.01	8.57	6	5.5	7.2	6.8	6.9

20.4.7: Question 7

A farmer exposed 7 pairs of chickens to dim and higher light conditions and recorded two aspects of laying under each condition. The H_0's are that the two variables (floor laying and nest laying of the chickens) were the same in dim and higher light conditions, then it is appropriate to use a test for comparison.

Table 20.27: Laying of chickens under different lights.

Pair	Floor-Dim	Nest-dim	Floor-high	Nest-high
1	300	295	80	60
2	240	260	120	140
3	250	280	170	160
4	220	250	90	120
5	160	160	150	180
6	170	150	110	90
7	300	290	260	120

20.4.8: Question 8

The weights at autopsy of 25 chickens suffering from a certain disease were as follows:

All the knowledge belongs to God.

Table 20.28: The weights of the chickens.

Weight (grams)				
859	1073	1041	1166	1117
962	1051	1064	1141	1202
973	1001	1016	1168	1255
904	1012	1002	1146	1233
920	1039	1086	1140	1348

Can we conclude from these data that the sampled population is not normally distributed with a mean of 1050 and a standard deviation of 50? Determine the p value for this test.

20.4.9: Solution

- Question 1: GLSD.
- Question 2: RCBD or factorial ANOVA (two-way).
- Question 3: One-way repeated measures ANOVA.
- Question 4: CRD.
- Question 5: RCBD under factorial or two-way ANOVA.
- Question 6: Independent samples T test.
- Question 7: Two-way repeated measures ANOVA.
- Question 8: Test of normality and one sample K-S test.

20.5: EXERCISE 5

20.5.1: Question 1

An experiment was set up to study genetic parameters for chicken. The full-sib families from a half-diallele cross with five parents represent the treatments which are evaluated for several growth traits, such as weight and breast yield, in a field trial at two different sites. More specifically, the basic arrangement consists of 6 blocks per site, each block containing 10 pens to which the ten full-sib families are assigned at random. In each pen 5 chickens from the assigned full-sib family are kept. Both sites are located in the Islamabad, but one site has colder climate than the other, being thus more prone to cold ventilation.

Formulate the design.

20.5.2: Question 2

Recognize the following design.

Table 20.29: Effect of molasses dose and types of mixing on feed composition.

Type of mixing	Form of active matter						
	0.5%	1%	1.5%	2%	2.5%	3%	3.5%
1	A 98	B 117	C 89	D 64	E 63	F 123	G 244
2	B 69	E 67	A 70	G 70	F 111	D 60	C 218
3	C 37	F 83	G 83	B 74	D 70	A 75	E 169
4	D 65	G 60	E 91	F 56	C 61	B 59	A 150
5	E 56	D 44	B 70	C 68	A 88	G 111	F 220

All the knowledge belongs to God.

20.5.3: Question 3

A pharmaceutical company conducts an experiment to compare 5 drugs. Thirty chickens are available for the trial. Each drug is injected into 6 randomly selected chickens. All the chickens are very similar. After an appropriate period 2 blood samples are taken from each animal and duplicate analyses are made for each blood sample.

Formulate the design.

20.5.4: Question 4

An experiment was conducted to investigate 4 feeds (F) and a vitamin (V) supplement at 3 levels on fish yield. The following table gives the information conducted in 3 blocks (R1; R2; R3). Analyze the data and draw your conclusion:

Table 20.30: Effect of 4 feeds and a vitamin supplement on fish yield.

	R1			R2			R3		
Feed	**V1**	**V2**	**V3**	**V1**	**V2**	**V3**	**V1**	**V2**	**V3**
F1	25	26	28	24	27	27	25	28	29
F2	22	23	25	22	24	26	23	23	26
F3	25	26	27	26	27	29	25	26	28
F4	28	30	31	27	29	30	28	30	31

20.5.5: Question 5

The purpose of a study by a scientist was to investigate floor egg number at poultry farms with automatic nests and manual nests. During experiment, floor eggs were recorded in first two weeks from 500 hens. The data is shown in following table. Analyze the data. We are also interested in testing whether the assumption of the equality of variances can be assumed prior to performing a t-test.

Table 20.31: Number of floor eggs.

Auto nest	131	115	124	131	122	117	88	114	150	169
Manual nest	60	150	130	180	163	130	121	119	130	148

20.5.6: Question 6

A study was conducted on the lifestyle of DVM, Zoology, and Poultry students (male and female) of a university. A lifestyle inventory was administered on everyone who participated in the study. A high score on the test indicates a better lifestyle. Test scores (out of 15) are shown in following table. Analyze the data given in table and discuss your findings at 5% level.

Table 20.32: Data on lifestyle evaluation.

Group	**DVM**	**Zoology**	**BBA**
Male	3	7	11
	2	4	7
	2	7	9
	4	6	8
	2	6	10
Female	8	8	11
	4	10	9
	3	7	11

All the knowledge belongs to God.

	5	7	12
	5	8	12

20.5.7: Question 7

A university developed a training program for improving lab skills of students. In order to test its effectiveness, she conducted a study in which 11 students were randomly selected. They were tested for their skills before and after implementing the training program for 6 weeks. The data so obtained are shown in following table:

Table 20.33: Skill set before and after training program.

Post_test	16.4	20.4	14.8	13.9	13.4	12.8	13.9	13.8	15.3	15.4	16.8
Pre_test	13	13	10.4	11.7	16	10.1	12.2	12.1	14.1	12.7	10

20.5.8: Question 8

The data given below are results of 25 design points performed at five temperatures and with five different time periods, with the idea of establishing effects of the given factors on conversion at a poultry farm. To avoid inequality effects, five pens and five chickens were included in the experiment. So, 25 design points were done in five pens with five operators by design of experiment in such a way that each chicken used each pen only once at each temperature and for a constant conversion time period. Characters denote pens and numbers the chickens. Do the analysis of variance.

Table 20.34: Poultry farm experiment.

	Time (minutes)				
Temperature (F)	**30**	**60**	**90**	**120**	**150**
50	16 A1	40 B3	50 C5	20 D2	15 E4
65	30 B2	25 C4	62 D1	67 E3	30 A5
70	50 C3	50 D5	83 E2	85 A4	45 B1
75	80 D4	80 E1	95 A3	98 B5	70 C2
80	90 E5	92 A2	98 B4	100 C1	88 D4

20.5.9: Solution

- Question 1: Nested design.
- Question 2: LSD under factorial or three-way ANOVA.
- Question 3: CRD.
- Question 4: LSD under factorial or three-way ANOVA.
- Question 5: Independent samples T test.
- Question 6: Two-way repeated measure ANOVA.
- Question 7: Paired samples T test.
- Question 8: Factorial GLSD.

20.6: EXERCISE 6

20.6.1: Question 1

An experiment was conducted to determine the effect of diets A, B, C, D and a control E on liver cholesterol (mg/dl) in deer. Recognized sources of variation were body

All the knowledge belongs to God.

weight and age. The researcher randomly selected and randomly arranged the results of the experiment as following. Analyze it.

Table 20.35: Liver cholesterol.

Weight group	Age group				
	I	II	III	IV	V
1	A 3.3	B 3.3	D 3.0	C 3.2	E 2.4
2	D 3.7	E 2.8	B 3.3	A 3.4	C 3.3
3	C 3.5	D 3.5	A 3.6	E 2.6	B 3.5
4	E 3.3	A 3.7	C 3.7	B 3.8	D 3.4
5	B 3.4	C 3.9	E 3.2	D 3.7	A 3.6

20.6.2: Question 2

Meats, including those from chicken and turkey are used as poultry meat. A study assessed the juiciness of meat (out of 100) from free-roaming chicken and turkeys obtained from Rawalpindi. These values were also compared to those of beef and fish. We want to know if the juiciness is different among the four meat groups.

Table 20.36: Juiciness of different meats.

Meat Type						
Chicken		Turkey		Beef		Fish
26	14	37	37	11	15	44
28	16	56	25	29	27	76
29	25	51	23	20	22	4
26	37	62	13	10	34	55
10	45	4	42	39	35	58
21	25	39	35	32	32	74
14	22	38	10	38	37	11
32	33	40	27	36	27	139
19	31	58	41	16	44	69
30	26	61	23	27	13	94
10	32	49	49	17	33	48
35	8	64	30	56	9	37
36	25	82	50	28	32	66
25	29	38	87	20	37	72
33	37	39	68	25	34	87
33	21			21	27	26
18	18			31	22	71

20.6.3: Question 3

In a study of reaction time (minutes) under the influence of xylazine (anesthetic), age is thought to be another factor that could affect the time. Test subjects (individuals) were classified into three age groups: 20-39, 40-59, 60 and over. In each age group each treatment (0mg, 0.1mg, 0.3mg per kg) was randomly assigned to 4 individuals. The following results were obtained (the reaction time is measured in minutes). Analyze data.

All the knowledge belongs to God.

Table 20.37: Influence of anesthetic due to age and dose.

Age	Dose of analgesic		
	0mg	**0.1mg**	**0.3mg**
20-39	.42	.49	.65
	.45	.47	.60
	.39	.46	.70
	.40	.51	.66
40-59	.51	.70	1.05
	.55	.69	1.10
	.53	.73	.98
	.50	.75	1.12
60 and over	.60	.85	1.25
	.59	.79	1.20
	.58	.88	1.30
	.61	.90	1.27

20.6.4: Question 4

In a study, test score of 9 males and 9 females was studied by quiz. Can it be concluded from the data that the scores of males and females were different at 0.05 significance level? The normality assumption has been violated.

Table 20.38: Test scores of students.

Female	13	13	12	9	9	15	15	15	17
Male	11	11	8	8	9	8	8	8	8

20.6.5: Question 5

Each person who came to the clinic was screened for dehydration level. Those who were diagnosed as "moderately depressed" were invited to participate in a treatment comparison study we were conducting. The IV is whether patients received cognitive-behavioral therapy or a 'support group control'. Because of ethical concerns, patients were not randomly assigned to treatment conditions. Rather, each was permitted to choose which treatment they would receive. In addition, each patient was permitted to decide when they would begin treatment. The DV score (rating of depression – bigger scores – more depression = "poorer") after 18 weeks of therapy was the output.

Formulate the design.

20.6.6: Question 6

In studying the effects of pollution, seedlings are usually exposed to specified pollutants for a certain period of time (say 6 hours) during the day for several weeks after which an evaluation is made. The seedlings are put in a pollution chamber which then receives the pollutant. Consider a specific case in which an investigator wants to compare 4 pollutants: P_0 = filtered air, $P_1 = O_3$, $P_2 = +NO_2$, and $P_3 = O_3 + NO_2$ at specified levels of concentration. He has only limited resources. In particular, he has only 8 pollution chambers in each of which he can put 3 seedlings. He feels that this is not adequate. So he decides to use the same chambers for 6 hours during the day and for 6 hours during the night, that is, 48 seedlings are used for the experiment. He is sure that there will be systematic differences between the day and night results. Recognize the design.

All the knowledge belongs to God.

20.6.7: Question 7

While developing weight training schedule, a coach wanted to know whether back strength differs among the university athletes playing soccer, wrestling, and hockey. He organized an experiment in which 10 soccer players, 14 wrestlers, and 12 hockey players were selected. Back strength for these subjects was tested by using the leg dynamometer. The data so obtained is shown in the following table. Analyze the data.

Table 20.39: Data on back strength in kg.

Soccer	68	88	100	116	72	66	88	78	90	92				
Wrestling	122	88	94	100	110	120	88	82	104	110	98	105	102	95
Hockey	78	82	96	110	120	88	92	98	88	110	88	86		

20.6.8: Question 8

Three different cooked fishes were evaluated by a panel of eight poultry scientists (judges). The judges are considered a random panel of all possible judges. The fishes are evaluated on a 100-point scale and were presented in random order to each judge. The following results obtained:

Table 20.40: Scores of judges.

	Fish score		
Judge	**1**	**2**	**3**
1	85	88	93
2	90	89	94
3	88	90	98
4	91	93	96
5	92	92	95
6	89	90	95
7	90	91	97
8	91	89	98

Recognize the design and analyze data.

20.6.9: Solution

- Question 1: LSD.
- Question 2: CRD or one-way ANOVA.
- Question 3: RCBD under factorial.
- Question 4: Mann-Whitney U test.
- Question 5: Independent samples T test.
- Question 6: Two-way ANOVA.
- Question 7: One-way ANOVA.
- Question 8: Two-way repeated measure ANOVA.

20.7: EXERCISE 7

20.7.1: Question 1

The dean of university wants to see whether there is a significant difference in ages of DVM and Zoology students. He selects a sample of defined students from each group. The ages are shown in Table 20.41. At $\alpha = 0.05$, decide if there is enough evidence to reject the claim of no difference in the ages of the two groups.

All the knowledge belongs to God.

Table 20.41: Ages of students.

DVM	22	20	19	23	26	28	26
	25	18	23	24	27	26	18
	18	30	22	18	18	19	32
	19	26	35	29	19	22	18
	26	23	21	21	23	18	20
	22	21	18	21	21	22	18
	19	25					
Zoology	18	20	19	18	22	25	24
	23	18	23	22	28	25	20
	26	30	22	22	22	21	18
	19	26	35	19	19	18	19
	29	23	21	19	36	27	27
	20	21	18	19	23	20	19
	20	25					

20.7.2: Question 2

In a university, a researcher wanted to investigate whether IQ of DVM and BBA students differs due to the very nature of their studies. He conducted a study in which 15 DVM and 15 BBA students who represented the university in the competition were selected for the study. The IQ was tested on them before the competition, which is shown in following table. The assumption of normality is violated. Analyze the data.

Table 20.42: IQ of DVM and Poultry students.

BBA	38	40	44	27	60	55	34	56	54	60	37	60	60	36	60
DVM	45	44	41	39	80	65	45	50	65	65	45	72	65	42	58

20.7.3: Question 3

Patients (adult chicken) suffering from 'calcium deficiency' often suffer critical losses in bone mineral density (BMD). Calcium gluconate is one medication prescribed to build or prevent further loss of BMD. A veterinarian looked at patients giving calcium gluconate to determine if a difference existed in the mean percent change in BMD among five different primary diagnosis classifications. Group 1 patients were diagnosed with weak shell eggs (WSE). Group 2 patients had lameness. Group three patients had arthritis. Group 4 patients had cage layer fatigue (CLF) and group 5 patients had osteoporosis with no other disease identified in the medical record. Changes in BMD are shown in the following table:

All the knowledge belongs to God.

Table 20.43: Mean % change in BMD.

WSE	CLF	Lameness	Arthritis	Osteoporosis
11	7	3	-4	11	3
24	6	0.2	-8	-1	15
10	5	8	4	4	5
-3	-4	3	-0.09	7	2
7	5	-1.3	-0.1	4	6
3.3	2	11	1.3	1.2	20
1.4	-3.9	3.9	5.2	-2.8	3.2
-1.8	9.6	7	10	4	9
5	5	6	1	4	6
4	1		0.4	1	26
6	7		0.5	-0.2	
-5	0.07		4	5	

20.7.4: Question 4

The nutritive value of a certain canola meal was measured taking different specimens representing six specimens of two varieties grown in two randomly taken regions. The results are as follows.

Analyze the data.

Table 20.44: The nutritive value of canola meal (mg).

Regions	Varieties	
	A	B
I	7	13
	12	14
	6	13
	9	13
	9	12
	6	14
II	9	9
	9	13
	5	13
	8	9
	8	10
	5	10

20.7.5: Question 5

A researcher wanted to compare the effects of three feed additives A, B and C on yield of broiler. The birds had first to be individually fed in house. But only three pens were available. Since three additives were to be tested, and only three pens were available, he decided to use a design with rows representing feed periods, and columns representing chickens assigned in each pen. To obtain sufficient degree of freedom for error and sufficient replications for reasonable precision, he wanted to run four squares. The experimental plan and yield during the last five days of each 42-day period are given below. Recognize the design.

All the knowledge belongs to God.

Table 20.45: Effects of three feed additives on yield of broilers.

Square	Period	11	12	13
I	11	C 115	A 139	B 127
	12	A 138	B 209	C 224
	13	B 125	C 186	A 172
II	21	21	22	23
	22	C 176	B 163	A 135
	23	A 186	C 201	B 175
		B 146	A 101	C 134
III	31	31	32	33
	32	C 186	B 194	A 166
	33	A 130	C 180	B 162
		B 123	A 97	C 137
IV	41	41	42	43
	42	A 128	C 154	B 150
	43	C 137	B 129	A 106
		B 164	A 138	C 168

20.7.6: Question 6

Cooling Pad is made from cellulose fiber and paper. An experiment is conducted to study the effect of using cellulose fiber along with paper. The researchers make 18 pads by varying the amount of paper (50%, 30% and 20%), the amount of fiber (50%, 70%, or 80%%), and the width of pad (4 inch, 5 inch and 6 inch). The response is the actual density of the pads produced. Analyze the data.

Table 20.46: Density of pads.

	Paper %			Width (inch)		
Fiber %	50%	30%	20%	4	5	6
50	41	42	42	44	46	48
70	43	44	45	48	48	51
80	45	46	46	50	50	50

20.7.7: Question 7

A feed scientist developed a new feed by altering part of the pelleting process. One outcome was the yield of turkey, measured in kgs. Ten custom made feed bags were offered, and yield was measured for yield after 7 days and again after 35 days. The results are shown in the following table. Note that there is considerable variation from feed to feed. For example, feed 5 had yield on both days, but feed 7 had high readings both times. Analyze the data.

Table 20.47: Yield of turkeys, measured in kgs.

Feed	Day 7 (Y1)	Day 30 (Y2)
1	7296	5544
2	6325	6120
3	8003	5720
4	5013	2508
5	4637	3743
6	8525	5272
7	9445	8189
8	8794	6794
9	5213	4409
10	3399	4083

20.7.8: Question 8

A researcher wanted to assess the effects of Diet and Exercises on weight gain in 10 deer. The participants were enrolled in four trials:

a) no diet and no exercises
b) diet only
c) exercises only
d) diet and exercises combined.

Each participant performed all four trials. The order of the trials was counter-balanced and sufficient time was allowed between trials to allow any effects of previous trials to have dissipated (i.e., a "wash out" period). Each trial lasted nine weeks and the weight loss score was measured at the beginning of each trial (t1); at the midpoint of each trial (t2); at the end of each trial (t3).

Recognize the design.

20.7.9: Solution

- Question 1: Independent samples T test.
- Question 2: Mann-Whitney U test.
- Question 3: Design name: CRD or one-way ANOVA.
- Question 4: Factorial ANOVA (two-way).
- Question 5: Three-way ANOVA.
- Question 6: LSD or Three-way ANOVA.
- Question 7: Repeated measures ANOVA.
- Question 8: Repeated measures crossover.

20.8: EXERCISE 8

20.8.1: Question 1

A fish farmer investigated the effect of mechanical stress in ponds on the growth of fish. Individually reared fishes were divided into two groups. Those in the first group were stressed by shaking water pond for 20 minutes twice daily, while those in the second group (the control group) were not shaken. After 6 weeks of growth the fishes were harvested, and total yield was measured for each fish. Analyze the data given in the following table.

All the knowledge belongs to God.

Table 20.48: Yield of fish.

Control	314	320	310	340	299	268	345	271	285
Stressed	283	312	291	259	216	201	267	326	241

20.8.2: Question 2

A scientist investigated the effectiveness on egg cessation of a hormone, high salt feed, or high fiber feed. Forty-eight breeders were assigned to one of three treatments. At their first egg cessation, the number of hours were noted. The results are shown in the following table:

Table 20.49: Egg cessation in breeders.

Egg cessation (hours)					
Hormone	High salt feed			High fiber feed	
15	8	60	90	8	80
17	10	60	90	15	80
18	15	60	90	25	82
20	20	60	95	25	86
20	22	60	96	25	87
20	24	60	98	26	90
30	25	60	98	30	90
37	26	66	99	34	90

Recognize the design.

20.8.3: Question 3

A researcher conducted studies to test the effects of flushers: furosemide and sugar solution (16%) on ascites. The 10 ostriches, ages 22–40 weeks, were chosen for study. Below is the urine output volumes (ml) following ingestion of furosemide, sugar solution and water (placebo):

Table 20.50: Urine output volumes (ml).

Subject No.	1	2	3	4	5	6	7	8	9	10
Furosemide	141	610	117	114	515	58	43	114	720	275
Sugar solution	131	600	217	214	315	100	49	120	550	365
Control (placebo)	237	161	160	149	147	144	88	118	144	890

Recognize the test.

20.8.4: Question 4

Abuse of insecticide at farm can produce various environmental contaminations. In an investigation of the mechanism of these toxic effects, researchers measured the concentrations of various chemicals in the litter that had been exposed to a pyrethrins, and also in unexposed litter. The concentrations of the litter pyrethrins in six pens (exposed to these chemicals) and five control pens, are given in following table:

All the knowledge belongs to God.

Table 20.51: Pyrethrins concentration in litter (mg/g).

Pyrethrin (Group 1)	543	523	431	635	564	549
Control (Group 2)	535	385	502	412	387	

20.8.5: Question 5

Consider an experiment comparing different feed managements for weight gain of broilers. Combinations of different types and amounts of feeds are assigned to individual bird in each of several selected rows of chickens. The chickens in each row are quite uniform, but there exist row-to-row differences due to environmental conditions. For purposes of inference several varieties of chickens were included in the experiment. Weight gain is checked at the end of each week till 5 weeks.

Formulate the design.

20.8.6: Question 6

Table 20.52: Different growth rate of chicken under diet and strains.

Exercise (hours)	Diet	Growth 1	Growth 2	Growth 3
Ross	A	112	166	215
		111	166	225
		89	132	189
Cobb	B	95	134	186
		66	109	150
		69	119	177
Indian River	C	125	177	241
		85	117	186
		97	137	185
Hubbard	D	93	151	217
		77	122	178
		78	119	173
Arber Acres	E	81	134	205
		88	133	180
		88	157	224
Desi	F	58	99	131
		85	132	186
		78	110	164

Recognize the design.

20.8.7: Question 7

In an experiment, three groups of chickens were given three different feeds for 6 weeks to see their effectiveness on weight gain. A control group was also taken in the study on which no feeding was imparted. The performance was measured before and after the treatment for 6 weeks, and the data so obtained is shown in following table. Apply test to find as to which feed is the best in improving yield among the chickens. Test your hypothesis at 0.05 and 0.01 level of significance.

Table 20.53: Data on vertical jump performance (in inches) in different depth jump.

Ross	Cobb	Hubbard	Control

All the knowledge belongs to God.

	Pre	Post	Pre	Post	Pre	Post	Pre	Post
SN	45	47	45	47	53	56	53	52
1	38	40	43	45	52	55	45	44
2	60	63	44	48	36	40	47	48
3	41	43	47	49	43	48	54	54
4	46	48	43	46	49	53	40	40
5	49	51	40	41	51	54	39	43

20.8.8: Question 8

Table 20.54: Four treatments applied on subjects.

Period	**Subjects**				**Subjects**					**Subjects**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	4	3	2	1	4	3	2	1	4	3	2	1
2	1	2	3	4	3	4	1	2	2	1	4	3
3	2	1	4	3	1	2	3	4	3	4	1	2
4	3	4	1	2	2	1	4	3	1	2	3	4

Recognize the design.

20.8.9: Solution

- Question 1: Independent samples T test.
- Question 2: CRD.
- Question 3: two-way ANOVA.
- Question 4: Independent samples T test.
- Question 5: Repeated measures ANOVA.
- Question 6: Repeated measures ANOVA.
- Question 7: ANCOVA.
- Question 8: Crossover design.

All the knowledge belongs to God.

CHAPTER 21: PRACTICE EXERCISES

21.1: EXERCISE 1

21.1.1: Question 1

A researcher studied the flexibility of each of seven people, four of whom were labors and three of whom were sportsman. One measure she recorded was the "trunk flexion"—how far forward each of the man could stretch while seated on the floor. The measures (in centimeters) are shown in following table.

Analyze the data using appropriate test.

Table 21.1: Trunk flexion measurements.

Labor	Sportsman
38	48
45	59
58	61
64	

21.1.2: Question 2

Some researchers have observed a greater weight gain in Ross than in Cobb chickens. Suppose a study, conducted to compare the percent of weight gain due to feeding, yielded the following results.

Table 21.2: Percent of tracheobronchial retention of particles.

Percent weight gain		Percent weight gain	
Ross	Cobb	Ross	Cobb
60	47	57	54
12	13	63	14
56	33	29	19
75	55	66	46
12	21	25	30
30	28	40	36

Analyze the data and suggest a design.

21.1.3: Question 3

A researcher wants to do an experiment using chickens as the EUs. She comes to you for advice on how to set up the experiment. Here is the situation: She wants to compare the effects of feeding three levels of calcium, say 0, 1, 2%, on certain bone measurements (which can be observed only after the chickens have been sacrificed). She has available 4 groups of 6 chickens each, each group containing 3 females and 3 males. Each group comes from a different breed of mice.

Formulate a design.

21.1.4: Question 4

An experiment was conducted to investigate the effects of four different diets D1, D2, D3, and D4 on daily gains (g) in weight of six chicks of 6 weeks old. The following data are related to gain in weights. Analyze the data and test whether all the four diet treatments are equally efficient and if not, which diet is the best.

All the knowledge belongs to God.

Table 21.3: Weight gain data under four diets.

D1	D2	D3	D4
30	32	34	36
27	28	25	30
23	27	26	28
25	29	27	33
23	26	25	32
26	29	29	30

21.1.5: Question 5

A teacher examined students using different teaching methodologies. In this study, 18 of the students completed an exam at baseline, and after 1, 3 and 6 weeks. Following table shows the data for these subjects who appeared in exam (out of 100). The goal of the experiment was to determine if subjects would report improvement over time. We wish to know if there is a difference in the mean test scores among the four points in time.

Table 21.4: Test scores at four different times.

Subject	Baseline	Month 1	Month 3	Month 6
1	80	60	95	100
2	95	90	95	95
3	65	55	50	45
4	50	45	70	70
5	60	75	80	85
6	70	70	75	70
7	80	80	85	80
8	70	60	75	65

21.1.6: Question 6

We illustrate the procedure of a farm trial on chicken. The trial tested varying numbers of feeds. The main character of interest was on shank length (cm). Shank length, measured before the chickens were placed on the plant, is known to affect the length later. Data on shank length after treatment (Y) and initial weight (X), for each of the 20 experimental chickens (5 treatments and 4 replications), are shown in the following table. Analyze the data.

Table 21.5: Shank length in chickens.

Treatment	Rep.1		Rep.2		Rep.3		Rep.4	
Feed	X	Y	X	Y	X	Y	X	Y
1	5	5	12	12	11	11	5	8
2	7	7	9	9	14	8	9	8
3	9	9	5	5	12	13	5	7
4	7	6	10	10	6	8	8	8
5	8	8	5	5	13	11	5	5

21.1.7: Question 7

In a 9 students PhD class, extra coaching classes was arranged to improve the stat score, given below the score of 9 students before and after coaching. Do all appropriate steps other than analysis?

All the knowledge belongs to God.

Table 21.6: Stat scores before and after coaching classes.

No.	1	2	3	4	5	6	7	8	9
Before	60	75	55	25	14	45	65	85	47
After	72	80	50	35	25	45	65	80	60

21.1.8: Question 8

In this experiment dogs were given drugs on request. Three drugs were compared in the experiment and each dog received two different drugs. The allocation of drugs to patients at the first request was random within overall restrictions of approximate equality of drug replication. The allocation of the drug at the second request, the second drug being not the same as the first drug, was similarly random. The full results of the trial are shown in following table. Analyze the data.

Table 21.7: Pain relief under two drugs.

		Patients							
Period	**Drug**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	T1	2	6	4	13	5	8	4	
2	T2	10	8	4	0	5	12	4	
1	T2	2	0	3	3	0			
2	T1	8	8	14	11	6			
1	T1	6	7	6	8	12	4	4	
2	T3	6	3	0	11	13	13	14	
1	T3	6	4	4	0	1	8	2	8
2	T1	14	4	13	9	6	12	6	12
1	T3	12	1	5	2	1	4	6	5
2	T2	11	7	12	3	7	5	6	3
1	T2	0	8	1	4	2	2	1	3
2	T3	8	7	10	3	12	0	12	5

21.2: EXERCISE 2

21.2.1: Question 1

Researchers were interested in the effect that feed has on the growth rate of fish. They created three treatment groups in an experiment: low protein, high protein, and control. The response variable in their experiment was the weight of the chickens in a farm after 35 days of growth. They had 5 pens for each of the 3 treatments, for a total of 15 observations. However, the pens were arranged near cooling pad and they wanted to account for the effect of differing amounts of cooling. Thus, they created 5 blocks—each block was a fixed distance away from the pad (block 1 being the closest through block 5, the farthest). Within each block the three treatments were randomly assigned. Recognize the design with its components.

All the knowledge belongs to God.

Table 21.8: Organization of blocks for chickens.

Cooling Pad	Block 1	Block 2	Block 3	Block 4	Block 5
	high	control	control	control	high
Feed Protein	control	low	high	low	low
	low	high	low	high	control

21.2.2: Question 2

An experiment was conducted to determine the requirements for protein and the amino acid threonine. Specifically, this experiment will examine combinations of the amount of protein in diet and the amount of threonine in diet. The general protein in the diet is threonine deficient. There are 8 levels of threonine (0.2 through 0.9% of diet) and 5 levels of protein (8.68, 12, 15, 18, and 21% of diet), for a total of 40 treatments.

Two-hundred chickens were used. On the second day after arrival, all chickens were weighed, and they were separated into 5 groups of 40 to provide groupings of approximately uniform weight. The 40 chickens in each group were randomly assigned to the 40 treatments. Body weight and feed consumption were measured twice weekly, and the response we consider is average daily weight gain over 21 days.

Recognize the design.

21.2.3: Question 3

Thirty-two animals were fed with four different feeds. The following figures give the gain in body weight after 2 months. Analyze the data and draw your conclusion: whether all the four feeds are equally efficient and which feed is the best feed:

Table 21.9: Weight gain effect due to four diets.

Feed 1	Feed 2	Feed 3	Feed 4
12	14	12	14
12	14	12	14
12	14	12	15
13	14	13	15
14	14	12	15
12	13	13	14
13	14	13	15
13		12	15
13		12	
13			

21.2.4: Question 4

A scientist wanted to show the effectiveness of a central four-quadrant sleeve and screw in anterior cruciate ligament reconstruction. The researchers performed a series of reconstructions on cadaveric knees. The following table shows the loads (in newtons) required to achieve different graft laxities (mm) for seven specimens (data not available for one specimen) using five different load weights. Graft laxity is the separation (in mm) of the femur and the tibia at the points of graft fixation. Is there sufficient evidence to conclude that different loads are required to produce different levels of graft laxity? Let α=0.05.

All the knowledge belongs to God.

Table 21.10: Loads (in newtons) required to achieve different graft laxities (mm).

Specimen	Graft Laxity (mm)				
	1	2	3	4	5
1	297	297	297	297	297
2	264	304	336	358	379
3	188	188	188	188	188
4	159	194	211	222	228
5	228	282	282	334	334
6	100	105	106	107	108
7	116	140	182	209	215

21.2.5: Question 5

A farmer is studying the burning rate of brooder from three production processes. Four batches of brooders are randomly selected from the output of each process, and three determinations of burning rate are made on each batch. The results are as following. Analyze the data and draw conclusions.

Table 21.11: Burning rate of brooders.

Batch	Process-1				Process-2				Process-3			
	1	2	3	4	1	2	3	4	1	2	3	4
	25	19	15	15	19	23	18	35	14	35	38	25
	30	28	17	1	17	24	21	27	15	21	54	29
	26	20	14	13	14	21	17	25	20	24	50	33

21.2.6: Question 6

A teacher examined test score before and after teaching a statistics course in students with biology background. He measured score by calculating the performance in presentation before and after teaching. The goal of teaching is to increase research skills, which is measured as a percent. The data are shown in following table. We wish to know if these data provide sufficient evidence to allow us to conclude that teaching increases research skills.

Table 21.12: test score in students with presentations of teaching.

Pre-op (%)	22	63	96	9	3	50	33	69	64	18	0	34
Post-op (%)	63	91	59	37	10	19	41	87	86	55	88	40

21.2.7: Question 7

Fish feed was stored using four different methods, and its nutritional value (kcal) was measured. Are there significant differences among different storage methods?

All the knowledge belongs to God.

Table 21.13: Nutritional values of feed under four storage conditions.

TR1	TR2	TR3	TR4
14	18	18	22
13	16	19	18
14	18	18	17
11	17	19	18
12	13	14	19
15		17	19
16			20

21.2.8: Question 8

A poultry scientist wanted to evaluate the effect of feed (Islamabad feed and Hi-tech feed) and chicken (Ross, Cobb and Hubbard) on broiler yield during a trial. Broilers were randomly divided into three groups. These groups were randomly assigned the feed treatment. After the treatment, their weight gain percentage was recorded by means of scale. The data so obtained are shown in following table.

Recognize the design.

Table 21.14: Data on yield percentage after the feed treatments.

Group	Ross	Cobb	Hubbard
Islamabad feed	20	8	27
	24	10	26
	21	11	31
	16	6	24
	21	9	26
Hi-tech feed	23	15	15
	20	18	19
	26	12	20
	19	10	12
	20	13	18

21.3: EXERCISE 3

21.3.1: Question 1

Choline and selenium (Se) supplementation of poultry feed are common practice. To better understand the effects of feed supplementation on the cellular retention of choline, researchers conducted an experiment by fortifying feed with low and high levels of Se and choline. The feeds were digested in a simulated gastrointestinal tract, and cellular choline and Se retention were measured (mg Se/mg cell protein).

Recognize the design.

21.3.2: Question 2

A scientist studied the effects of reminiscence therapy for zoology students with depression. She studied 12 students 22 years or older residing for 3 months or longer in hostel. For this study, depression was measured by the Geriatric Depression Scale (GDS). Higher scores indicate more severe depression symptoms. The participants received reminiscence therapy for long-term care, which uses family photographs, scrapbooks, and personal memorabilia to stimulate memory and conversation among group members. Pre-treatment and post-treatment depression scores are given in the following table. Can we conclude, based on these data, that subjects who

All the knowledge belongs to God.

participate in reminiscence therapy experience, on average, a decline in GDS depression scores? Let α=0.01

Table 21.15: Depression score measured by the Geriatric Depression Scale.

Pre-GDS	12	10	16	2	12	18	11	16	16	10	14	21
Post-GDS	11	10	11	3	9	13	8	14	16	10	12	22

21.3.3: Question 3

A scientist designed a study to evaluate selenium intake in broiler breeders. The researchers studied a cohort of 6 breeders for three consecutive weeks. One of the outcome variables was the selenium intake per day. The researchers examined dietary journals of the subjects over the course of 2 weeks and then computed the average daily selenium intake. The following table shows the average daily selenium intake values (ug/day) for the 6 breeders in weeks 1, 2, and 3 of the study. Analyze the data.

Table 21.16: Average daily selenium intake values (ug/day) for the 16 women.

Subject	Week 1	Week 2	Week 3
1	112	121	94
2	106	121	145
3	102	121	130
4	103	90	135
5	112	121	145
6	106	98	145

21.3.4: Question 4

Four different breeds of fishes were treated with five vitamins for improving weight. Type of breed was criteria for assigning fishes in four different blocks. Each block is assigned with four fishes. The effect of these vitamins on weight was tested by weekly weight (grams) after treatment.

Analyze the data.

Table 21.17: Weekly weight (grams) after vitamin treatment.

Vitamin	Breed 1	Breed 2	Breed 3	Breed 4
V1	42	54	72	88
V2	44	57	76	92
V3	45	52	78	86
V4	42	60	73	93
V5	41	61	78	92
V6	46	65	82	99

21.3.5: Question 5

The yield of a farm process is being studied. The two factors of interest are temperature and ventilation (cfm). Three levels of each factor are selected; however, only nine runs can be made in one day. The experimenter runs a complete replicate of the design on each day. The data are shown in the following table. Analyze the data, assuming that the days are blocks.

All the knowledge belongs to God.

Table 21.18: Yield under different temperature and pressures.

	Day 1 ventilation			Day 2 ventilation		
Temperature	**1.5**	**2.5**	**3.0**	**1**	**2.5**	**3**
Low	86	84	86	86	85	87
Medium	88	87	89	89	90	90
High	89	90	91	92	93	94

21.3.6: Question 6

A manager is studying the effectiveness of feed delivery methods to the farm. Three different types of hand trucks have been developed, and an experiment is performed in the company's laboratory. The variable of interest is the delivery time in minutes (y); however, delivery time is also strongly related to the case volume delivered (x). Each hand truck is used four times and the data that follow are obtained. Analyze these data and draw appropriate conclusions. Use alpha=0.05.

Table 21.19: Delivery time under 3 hand trucks.

Hand truck type					
1		2		3	
y	X	y	x	y	x
27	24	25	26	40	38
44	40	35	32	22	26
33	35	46	42	53	50
41	40	26	25	18	20

21.3.7: Question 7

IQs sample of 30 people arrested for drug abuse in a Islamabad were as follows:

Table 21.20: IQs of a sample of adolescents arrested for drug abuse.

IQ					
95	100	91	106	109	110
98	104	97	100	107	119
92	106	103	106	105	112
101	91	105	102	101	110
101	95	102	104	107	118

Do these data provide sufficient evidence that the sampled population of IQ scores is not normally distributed with a mean of 105 and a standard deviation of 10? Determine the p value.

21.3.8: Question 8

In a study of pulmonary effects on guinea pigs, a scientist exposed ovalbumin (OA)-sensitized guinea pigs to regular air, benzaldehyde, or acetaldehyde. At the end of exposure, the guinea pigs were anesthetized and allergic responses were assessed in bronchoalveolar lavage (BAL). One of the outcome variables examined was the count of eosinophil cells, a type of white blood cell that can increase with allergies. Table gives the eosinophil cell count ($\times 10^6$) for the three treatment groups.

Can we conclude that the three populations represented by the three samples differ with respect to eosinophil cell count? We can so conclude if we can reject the null hypothesis that the three populations do not differ in eosinophil cell count. Assumptions are seriously violated.

All the knowledge belongs to God.

Table 21.21: Eosinophil Count for Ovalbumin-Sensitized Guinea Pigs.

Eosinophil cell count ($\times 10^6$)		
Air	**Benzaldehyde**	**Acetaldehyde**
12.2	3.6	54.3
28.4	4	27.8
28.1	6.4	66.8
38.6	21.1	46.2
54.9	3.3	30.1

21.4: EXERCISE 4

21.4.1: Question 1

The purpose of an investigation by was to evaluate the analgesic effectiveness of a daily dose of oral methadone in patients with chronic neuropathic pain syndromes. The researchers used a visual analogue scale (0–100 mm, higher number indicates higher pain) ratings for maximum pain intensity over the course of the day. Each subject took either 20 mg of methadone or a placebo each day for 5 days. Subjects did not know which treatment they were taking. The following table gives the mean maximum pain intensity scores for the 5 days on methadone and the 5 days on placebo.

Do these data provide sufficient evidence, at the 0.05 level of significance, to indicate that in general the maximum pain intensity is lower on days when methadone is taken?

Table 21.22: Pain intensity scores for the 5 days.

Subject	1	2	3	4	5	6	7	8	9	10	11
Methadone	30	73	99	59	61	57	57	89	97	50	37
Placebo	57	70	98	62	67	70	68	96	98	63	64

21.4.2: Question 2

The purpose of a study was to examine the effect of voluntarily slowed respiration on the cardiac parasympathetic response to a threat (the anticipation of an electric shock). Subjects were 21 healthy college students whose mean age was 23 years with a standard deviation of 1.5 years. An equal number of subjects were randomly assigned to slow (six males, four females), fast (seven males, three females), and non-paced (five males, five females) breathing groups. Subjects in the slow- and fast-paced breathing groups regulated their breathing rate to 8 and 30 cpm, respectively. The nonpaced group breathed spontaneously. The following are the subjects' scores on the State Anxiety Score of State-Trait Anxiety Inventory after baseline and period of threat.

All the knowledge belongs to God.

Table 21.23: Scores on the State Anxiety Score of State-Trait Anxiety Inventory.

Slow paced		Fast paced		Non-paced	
Baseline	**Threat**	**Baseline**	**Threat**	**Baseline**	**Threat**
39	59	37	49	36	51
44	47	40	42	34	71
48	51	39	48	50	37
50	61	47	57	49	53
34	48	45	49	38	52
54	69	43	44	39	56
34	43	32	45	66	67

21.4.3: Question 3

A study was carried out by farmer to determine whether major differences in culling % in response to choline supplementation exist among widely-grown hybrids of broilers. The subplot treatments were five hybrids (H1, H2, H3, H4 and H5), and the main-plot treatments were choline doses of 0, 35, 70 and 105 mg/kg. The study was replicated two times. Culling data of five broiler hybrids grown in four levels of choline in a design experiment with two replications is given below. Analyze it.

Table 21.24: Grain yield of hybrids of maize in response to nitrogen fertilization.

		Choline dose (mg/kg)			
Replication	**Broiler hybrid**	**0**	**35**	**70**	**105**
	H1	130	150	170	165
	H2	125	150	160	165
I	H3	110	140	155	150
	H4	115	140	160	140
	H5	115	170	160	170
	H1	135	170	190	185
	H2	150	160	180	200
II	H3	135	155	165	175
	H4	130	150	175	170
	H5	145	180	195	200

21.4.4: Question 4

Suppose a dermatologist wants to study the effectiveness of 2 different preparations of a skin lotion using 2 different forms of application (for example, one vs. two applications per day). He has available 12 patients with a certain skin disease and he can apply one form of medication (that is, combination of preparation and frequency of application) to each arm of each patient. Even though the patients have the same disease, there exists considerable variation among them, but the two arms of a patient are quite homogeneous. Formulate suitable design.

21.4.5: Question 5

A scientist evaluated the effect of telephone follow-up on the physical well-being dimension of health-related quality of life in patients with cancer. One of the main outcome variables was measured by the physical well-being subscale of the Functional Assessment of Cancer Therapy Scale-General (FACT-G). A higher score indicates higher physical well-being. The following table shows the baseline FACT-G score and the follow-up score to evaluate the physical well-being during the 7 days after discharge from hospital to home for 20 patients who received a phone call 48–72 hours after discharge that gave patients the opportunity to discuss

All the knowledge belongs to God.

medications, problems, and advice. Is there sufficient evidence to indicate that quality of physical well-being significantly decreases in the first week of discharge among patients who receive a phone call? Let $\alpha=0.05$.

Table 21.25: FACT-G score and the follow-up score to evaluate the physical well-being.

Subject	Baseline	Follow-up	Subject	Baseline	Follow-up
1	16	19	11	25	14
2	26	19	12	21	17
3	13	9	13	14	22
4	20	23	14	23	22
5	22	25	15	19	16
6	21	20	16	19	15
7	20	10	17	18	23
8	15	20	18	20	21
9	25	22	19	18	11
10	20	18	20	20	22

21.4.6: Question 6

Suppose an industrial engineer is studying the hand insertion of electronic components on printed circuit boards to improve the speed of the assembly operation. He has designed three assembly fixtures and two workplace layouts that seem promising. Operators are required to perform the assembly, and it is decided to randomly select four operators for each fixture–layout combination. However, because the workplaces are in different locations within the plant, it is difficult to use the same four operators for each layout. Therefore, the four operators chosen for layout 1 are different individuals from the four operators chosen for layout 2. Because there are only three fixtures and two layouts, but the operators are chosen at random, this is a mixed model. The treatment combinations in this design are run in random order, and two replicates are obtained. The assembly times are measured in seconds and are shown in the following table.

Recognize the design.

Table 21.26: Assembly times of electronic components.

	Layout 1				Layout 2			
	1	2	3	4	1	2	3	4
Fixture 1	22	23	28	25	26	27	28	24
	24	24	29	23	28	25	25	23
Fixture 2	30	29	30	27	29	30	24	28
	27	28	32	25	28	27	23	30
Fixture 3	25	24	27	26	27	26	24	28
	21	22	25	23	25	24	27	27

21.4.7: Question 7

An experiment was conducted to see the effect of four treatments on ovulation (no. of eggs produced/fish in thousands) in rohu (Labeo rohita) fish. Four different categories of rohu fish was formed considering weight and treatment. The treatments were randomly assigned to the fish of different breeds. Test whether these treatments significantly differ or not:

All the knowledge belongs to God.

Table 21.27: Effect of four treatments on ovulation of four fish breeds.

	Cat 1	Cat 2	Cat 3	Cat 4
T1	80	100	140	180
T2	85	120	160	200
T3	82	125	170	190
T4	86	140	175	220

21.4.8: Question 8

Here are some data on a shell measurement (the length of the anterior adductor muscle scar, standardized by dividing by length; I'll call this "AAM length") in the mussel *Mytilus trossulus* from five locations: Tillamook, Oregon; Newport, Oregon; Petersburg, Alaska; Magadan, Russia; and Tvarminne, Finland, taken from a much larger data set used in a study. Analyze the data.

Table 21.28: AAM length of 5 regions.

Tillamook	Newport	Petersburg	Magadan	Tvarminne
0.057	0.087	0.103	0.070	
0.081	0.066	0.135	0.091	0.102
0.083	0.067	0.081	0.078	0.095
0.097	0.081	0.101	0.068	0.097
0.081	0.074	0.096	0.067	0.103
0.085	0.064	0.106	0.069	0.104
0.073	0.083	0.105	0.076	
0.065	0.072	0.068		
0.092				
0.083				

21.5: EXERCISE 5

21.5.1: Question 1

An experiment was conducted to determine the effect of three diet treatments (T1, T2, and T3) on daily gain body weight (g/d) of fish. Fishes of five different breeds were selected. In each breed, there were six animals to which each treatment was randomly assigned to two animals. Therefore, a total of 30 animals were used. Analyze the data and draw conclusions on whether:

- There exists significant difference between three different breeds of pigs with respect to weight gain.
- Three different treatments significantly differ with respect to weight gain.
- There exists any interaction effect between the breed and diet or not:

Table 21.29: Weight gain of fish breeds under 3 diets.

	Pig breed 1	Pig breed 2	Pig breed 3	Pig breed 4	Pig breed 5
T1	240	290	510	320	420
	250	275	520	340	410
T2	170	265	470	330	375
	180	260	480	300	380
T3	190	255	500	310	390
	210	265	490	290	395

All the knowledge belongs to God.

21.5.2: Question 2

A scientist investigated the effect of ultraviolet A (UVA) therapy used on chickens. One of the outcome variables is the egg production. The following table gives the production scores for 10 subjects measured at baseline and after eight treatments. Do these data provide sufficient evidence, at the 0.01 level of significance, to indicate that the combination therapy increases egg production scores?

Table 21.30: Scores under UVA therapy.

Subject	1	2	3	4	5	6	7	8	9	10
Baseline	6	8	13	17	6	15	7	5	10	12
After 8 treatments	5	12	5	4	0.4	4	1	3	4	5

21.5.3: Question 3

One of the prevention of COVID-19 is to wash hands with any soap again and again for 20 seconds. All of you are aware this statement given by WHO. Do you think this conclusion is arises after conducting a design of experiment? If yes plan that experiment among response, factors with their level, model and name that design along with its layout.

21.5.4: Question 4

The wing yield (grams) of three different varieties of chicken with application of three proteins is given below. Analyze the data to show whether:

Table 21.31: The yield (g) of 3 varieties of chicken.

Treatment	Variety 1	Variety 2	Variety 3
T1	9	9	12
	9	10	11
	10	11	11
T2	9	9	11
	9	8	12
	9	9	12
T3	12	12	15
	11	13	14
	11	13	15

21.5.5: Question 5

A researcher sampled a number of stations (between 4 and 7) on six streams known to be polluted by the heavy metals in the Rocky Mountain region of Colorado, USA. They recorded zinc concentration, and richness and species diversity of the diatom community and proportion of diatom cells that were the early successional species, *Achanthes minutissima*. The first analysis compares mean diatom species diversity (response variable) across the 4 zinc-levels groups (categorical predictor variable), zinc level treated as a fixed factor. The H_0 was no difference in mean diatom species diversity between zinc-level groups.

All the knowledge belongs to God.

Table 21.32: Specie diversity under 4 zinc levels.

ZINC	DIV	ZINC	DIV	ZINC	DIV	ZINC	DIV	ZINC	DIV
Back	2.27	Med	2.19	Low	1.83	Med	1.75	High	1.04
High	1.25	Med	2.1	Low	1.88	Low	2.83	Low	2.18
High	1.15	Back	2.2	Med	2.02	Back	1.53	Back	1.89
Med	1.62	Med	2.06	Med	1.94	Back	0.76	High	1.37
Back	1.7	High	1.9	Low	2.1	Med	0.8	Low	1.4
High	0.63	High	1.88	Low	2.3	Low	1.66	Back	1.98
Back	2.05	High	0.85	High	1.43	Med	0.98		

21.5.6: Question 6

Substances to be tested for cancer-causing potential are often painted on the skin of mice. The question arose whether mice might get an additional dose of the substance by licking or biting their cage-mates. To answer this question, the compound benzo(a)pyrene was applied to the backs of 10 mice: 5 were individually housed, and 5 were group housed in a single cage. After 48 hours, the concentration of the compound in the stomach tissue of each mouse was determined. Normality assumption is violated. The results (nmol/gm) were as follow:

Table 21.33: benzo(a)pyrene dose absorbed.

Singly housed	Grouped-housed
3.3	3.9
2.4	4.1
2.5	4.8
3.3	3.9
2.4	3.4

21.5.7: Question 7

Following information are pertaining to the gain in weight (g) per month of five groups of broilers administered with five different feeds. Analyze the data and test whether there exists any difference in effects of feeds on body weight of fishes or not. If yes, then find out the best feed:

Table 21.34: Gain in weight (g) of weight administered with 5 feeds.

Feed 1	Feed 2	Feed 3	Feed 4	Feed 5
109	94	160	110	75
104	87	155	125	78
111	81	135	117	70
117	81	142	18	75
105	95	155	120	80
135	105	155	132	80
142	105		135	55

21.5.8: Question 8

In a study, hemodynamic stresses were measured on subjects undergoing laparoscopic cholecystectomy. An outcome variable of interest was the ventricular end diastolic volume (LVEDV) measured in milliliters. A portion of the data appear in the following table. Baseline

All the knowledge belongs to God.

refs to a measurement taken 5 minutes after induction of anesthesia, and the term "5 minutes" refers to a measurement taken 5 minutes after baseline. The assumptions are violated.

Table 21.35: Hemodynamic stresses measured on subjects undergoing laparoscopic cholecystectomy.

	LEVDV (ml)	
Subject	Baseline	5 minutes
1	52	49
2	79	72
3	79	87
4	80	88
5	72	103
6	85	94
7	70	94
8	71	46
9	56	72
10	56	72

21.6: EXERCISE 6

21.6.1: Question 1

Calcium is an essential mineral that regulates the heart, is important for blood clotting and for building healthy bones. The National Osteoporosis Foundation recommends a daily calcium intake of 1000-1200 mg/day for adult men and women. While calcium is contained in some foods, most adults do not get enough calcium in their diets and take supplements. Unfortunately some of the supplements have side effects such as gastric distress, making them difficult for some patients to take on a regular basis.

A study is designed to test whether there is a difference in mean daily calcium intake in adults with normal bone density, adults with osteopenia (a low bone density which may lead to osteoporosis) and adults with osteoporosis. Adults 60 years of age with normal bone density, osteopenia and osteoporosis are selected at random from hospital records and invited to participate in the study. Each participant's daily calcium intake is measured based on reported food intake and supplements.

The data are shown below. Analyze the data.

All the knowledge belongs to God.

Table 21.36: Daily calcium intake of different patients.

Normal Bone Density	Osteopenia	Osteoporosis
1200	1000	890
1000	1100	650
980	700	1100
900	800	900
750	500	400
800	700	350
900	1100	400

21.6.2: Question 2

An enterprise examined 10 subjects with neck cancer and measured as one of the outcome variables an oral health condition score. Patients were randomly divided into two treatment groups. These were a placebo treatment (treatment 1) and an aloe juice group (treatment 2). Cancer health was measured at baseline and at the end of 2, 4, and 6 weeks of treatment. The goal was to discern if there was any change in oral health condition over the course of the experiment and to see if there were any differences between the two treatment conditions.

Table 21.37: Oral health condition scores at different points in times

Subject	Treatment 1=placebo 2-aloe juice	Tota c1	Total c2	Total c3	Total c4
1	1	6	6	6	7
2	1	9	6	10	9
3	1	7	9	17	19
4	1	6	7	9	3
5	1	6	7	16	13
6	2	6	11	11	14
7	2	6	7	6	6
8	2	12	11	12	9
9	2	5	7	13	12
10	2	6	7	7	7

21.6.3: Question 3

A researcher designed an experiment to assess the effects of prolonged inhalation of carbon monoxide. Fifteen laboratory chickens served as experimental subjects, while 10 similar birds served as controls. The variable of interest was hemoglobin level following the experiment. The results are shown in following table. We wish to know if we can conclude that prolonged inhalation of carbon monoxide reduces hemoglobin level. Assumptions are violated.

Table 21.38: Hemoglobin determinations (grams) for 25 laboratory chickens.

Exposed chickens (X)	Unexposed chickens (Y)

All the knowledge belongs to God.

14.4	17.4
14.2	16.2
13.8	17.1
16.5	17.5
14.1	15
16.6	16
15.9	16.9
15.6	15
14.1	16.3
15.3	16.8
15.7	
16.7	
13.7	
15.3	
14	

21.6.4: Question 4

A researcher conducted a study to determine if nursing students who were assigned to a home hospital (HH) experience differed from those traditionally placed (TP) in hospitals throughout their nursing training. A small subset of data is provided in the table below. In this data set, hospital placement is the between-subjects variable. Anxiety, as measured by Spielberger's State Anxiety Scale (where higher scores suggest higher levels of anxiety), is the within-subjects variable and is provided at four points in time during nursing training. Is there evidence that anxiety level changed through time for these nursing students? Is there a difference in anxiety between those in a home hospital placement versus traditional placement? Is there significant interaction between placement type and anxiety? Let $\alpha=0.05$.

Table 21.39: State Anxiety Scale.

Subject	Hospital placement	Anxiety 1	Anxiety 2	Anxiety 3	Anxiety 4
1	HH	51	33	12	31
2	HH	50	51	50	44
3	HH	65	58	45	37
4	HH	43	40	31	51
5	HH	67	56	50	42
6	TP	44	48	51	59
7	TP	44	50	54	40
8	TP	54	49	35	46
9	TP	38	38	32	37
10	TP	25	27	25	24

21.6.5: Question 5

A plant physiologist investigated the effect of flooding on root metabolism in two tree species: flood tolerant river birch and the intolerant European birch. Four seedlings of each species were flooded for one day, and four were used as controls. The concentration of adenosine triphosphate (ATP) in the roots of each plant was measured. The data (nmol ATP per mg tissue) are shown in the table, analyze it.

All the knowledge belongs to God.

Table 21.40: ATP per mg tissue of two plants.

River birch		European birch	
Flooded	**Control**	**Flooded**	**Control**
1.45	1.7	0.21	1.34
1.19	2.04	0.58	0.99
1.05	1.49	0.11	1.17
1.07	1.91	0.27	1.30

21.6.6: Question 6

Suppose we have 5 kinds of mice namely A, B, C, D and E. All are crossed with different aging criteria's, with respect to every means. A company wants to utilize their sense of smelling as they belongs to that class having long range detection smell system naturally. In future these mice's may use in collapsing a tall buildings in earthquake or any other disasters to trace the any survived body. For training purpose all selected kind of mice's gone through with three replicate. Company want to know the performance of mice's by kind and with no meal in first replicate, partial meal in second replicate and complete meal in third replicate. Identify the design of experiment.

Table 21.41: Sense of smelling of mice.

	No meal	**Partial meal**	**Complete meal**
Mice A	Smell range (low/high)	Smell range (low/high)	Smell range (low/high)
Mice B	Smell range (low/high)	Smell range (low/high)	Smell range (low/high)
Mice C	Smell range (low/high)	Smell range (low/high)	Smell range (low/high)
Mice D	Smell range (low/high)	Smell range (low/high)	Smell range (low/high)
Mice E	Smell range (low/high)	Smell range (low/high)	Smell range (low/high)

21.6.7: Question 7

One of the purposes of an investigation was to investigate the effect on CD4 T cell count of administration of intermittent interleukin (IL-2) in addition to highly active antiretroviral therapy (HAART). The following table shows the CD4 T cell count at baseline and then again after 12 months of HAART therapy with IL-2. Do the data show, at the 0.05 level, a significant change in CD4 T cell count?

Table 21.42: CD4 T cell count.

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**
CD4 T cell count at entry ($\times 10^6/L$)	173	58	103	181	105	301	169
CD4 T cell count at the end of follow-up ($\times 10^6/L$)	257	108	315	362	141	549	369

21.6.8: Question 8

The following data gives number of fruits per plant under four types of growth regulator treatments in 3 years. Analyze the data to show which growth regulator and which year have resulted maximum fruits per plant:

All the knowledge belongs to God.

Table 21.43: Fruits per plant under growth regulators in 3 years.

Method	Year 1	Year 2	Year 3
GR1	145	135	150
GR2	195	200	210
GR3	355	375	385
GR4	240	225	275

21.7: EXERCISE 7

21.7.1: Question 1

A completely randomized double-blind clinical trial was conducted to compare two drugs, ticrynafen (T) and hydrochlorothiazide (H), for effectiveness in treatment of high blood pressure. Each drug was given at either a low or a high dosage level for 6 weeks. The accompanying table shows the results for the drop (baseline minus final value) in systolic blood pressure (mm Hg):

Table 21.44: blood pressure (mm Hg) under two drugs.

Ticrynafen (T)		Hydrochlorothiazide (H)	
Low dose	High dose	Low dose	High dose
13.9	17.1	15.8	17.5
53	57	55	58

21.7.2: Question 2

Kindergarten students were the participants in a study conducted. The researchers studied the fine motor skills of 15 children receiving occupational therapy. They used an index of fine motor skills that measured hand use, eye–hand coordination, and manual dexterity before and after 7 months of occupational therapy. Higher values indicate stronger fine motor skills. The scores appear in the following table:

Table 21.45: Fine motor skills of 37 children receiving occupational therapy.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pre	91	61	85	88	94	112	109	79	109	115	59	85	112	85	88
Post	94	94	103	112	91	112	112	97	100	106	100	88	112	112	97

21.7.3: Question 3

The following data give the weight (q) of chicken for four different types of feeds in 3 different weeks. Analyze the data, and comment:

- Which type of feed is the best
- Which week has produced maximum weight per chicken
- Which type-year combination has produced highest yield per chicken

All the knowledge belongs to God.

Table 21.46: Weight (q) of chickens for 4 types of feeds in 3 weeks.

Method	Week 1	Week 2	Week 3
V1	5	5	5
	5	6	5
	4	6	5
V2	9	10	8
	8	10	7
	8	9	7
V3	4	6	4
	4	7	5
	4	6	5
V4	6	5	4
	6	5	4
	5	6	5

21.7.4: Question 4

A study investigated spinal canal dimensions in 30 subjects symptomatic with disc herniation selected for a discectomy and 45 asymptomatic individuals (control group). One of the areas of interest was determining if there is a difference between the two groups in the spinal canal cross-sectional area (cm2) between vertebrae L5/S1. The data in the following table are simulated to be consistent with the results reported in the paper. Do these simulated data provide evidence for us to conclude that a difference in the spinal canal cross-sectional area exists between a population of subjects with disc herniations and a population of those who do not have disc herniations? Let α=0.05.

Table 21.47: Spinal canal cross-sectional area (cm2) between vertebrae L5/S1.

Herniated disc group	2.6	2.5	1.9	3.2	3.5
	1.6	1.8	3.9	2.5	1.5
	2.3	2.6	3.5	2.2	2.8
	2.0	1.1	3	2.3	3.6
	2.0	3.7	2.4	2.5	2.1
	2.2	2.3	2.8	3.7	2.6
Control group	3.7	4.3	2.8	3.8	2.7
	1.3	2.3	3.6	1.6	3.5
	4.2	3	3.3	4	2.7
	3.1	3.9	4.3	3.7	2.2
	5	3.6	3	3.1	3.5
	5.4	3.3	2.6	3.7	4.3

21.7.5: Question 5

An experiment was conducted to study the effect of 4-week running and walking exercise on resting pulse rate. Three groups of subjects were selected for the study. Each group consisted of 5 subjects. The first group was given running, the second, the walking exercise, and the third served as control group. Resting pulse rate was measured before and after the treatment in all the three groups. The pre-exercise resting pulse rate might affect our response (post-exercise rate). The data so obtained are shown in table, analyze it.

All the knowledge belongs to God.

Table 21.48: Data on resting pulse rate (beat/min).

SN	Running Pre	Running Post	Walking Pre	Walking Post	Control group Pre	Control group Post
1	70	68	69	67	71	71
2	73	72	72	71	72	73
3	69	67	70	70	80	79
4	71	70	68	66	78	78
5	80	78	82	81	69	70

21.7.6: Question 6

We want to investigate the H_0 that the weight gain of deer is the same on each of four specified diets. Each deer is housed in a separate cage. A block consists of a group of four animals that we can be reasonably assured will experience identical environmental conditions (light, temp, draft, noise, etc.). Each block has each of its four animals assigned at random to one of the four experimental diets, so that each animal in a given block is to receive a different diet. The data (weight gains, in grams) are summarized in the following table:

Table 21.49: weight gain of deer under different diets and blocking.

Blocks	Diets 1	2	3	4
1	7	5.3	4.9	8.8
2	9.9	5.7	7.6	8.9
3	8.5	4.7	5.5	8.1
4	5.1	3.5	2.8	3.3
5	10.3	7.7	8.4	9.1

21.7.7: Question 7

Researchers studied the effect that exposure to ultraviolet-B radiation has on the survival of embryos of the western toad Bufo boreas. They conducted an experiment in which several B. borea embryos were placed at one of three water depths—10 cm, 50 cm, or 100 cm—and one of two radiation settings—exposed to UV-B radiation or shielded. The response variable was the percentage of embryos surviving to hatching. Formulate the layout.

21.7.8: Question 8

Nineteen chickens are assigned at random among 4 experimental groups. Each group is fed a different diet. The data are chicken body weights, in kilograms, after being raised on these diets. We wish to ask whether chicken weights are the same for all four diets.

Table 21.50: Chicken body weights under 4 feedings.

Feed 1	Feed 2	Feed 3	Feed 4
60	68	102	87
57	67	102	84
65	74	100	83
58	66	96	85
61	69		90

All the knowledge belongs to God.

21.8: EXERCISE 8

21.8.1: Question 1

A group of 24 broiler breeders was randomly divided into six subgroups of four and each broiler got selenium (Se) shot. Taking into account three levels of Se doses and two procedures of preparing Se, A and B, increase of fertility percent in breeders was measured sometime after injecting Se. The obtained results are shown in the following table, analyze it.

Table 21.51: Fertility percent in breeders.

		Doses		
		2.29	**3.63**	**5.57**
Preparation	**A**	17	64	62
		21	48	72
		49	34	61
		54	63	91
	B	33	41	56
		37	64	62
		40	34	57
		16	64	72

21.8.2: Question 2

For each of nine horses, a veterinary anatomist measured the density of nerve cells at specified sites in the intestine. The results for site I (mid region of jejunum) and site II (mesenteric region of jejunum) are given in the accompanying table. Each density value is the average of counts of nerve cells in five equal sections of tissue, but the normality assumption is violated. The null hypothesis of interest is that in the population of all horses there is no difference between the two sites.

Table 21.52: Nerve cell density at each of two sites.

Animal	**Site I**	**Site II**
1	50.6	38
2	39.2	18.6
3	35.2	23.2
4	17	19
5	11.2	6.6
6	14.2	16.4
7	24.2	14.4
8	37.4	37.6
9	35.2	24.4

21.8.3: Question 3

A dental research team wished to know if teaching people how to brush their teeth would be beneficial. Twelve pairs of patients seen in a dental clinic were obtained by carefully matching on such factors as age, sex, intelligence, and initial oral hygiene scores. One member of each pair received instruction on how to brush his or her teeth and on other oral hygiene matters. Six months later all 24 subjects were examined and assigned an oral hygiene score by a dental hygienist unaware of which subjects had received the instruction. A low score indicates a high level of oral hygiene. The assumption of normality is violated. The results are shown in following table:

All the knowledge belongs to God.

Table 21.53: Oral hygiene scores of subjects receiving oral hygiene instruction and not receiving instruction.

	Score	
Pair number	**Instructed (Xi)**	**Not instructed (Yi)**
1	1.5	2
2	2	2
3	3.5	4
4	3	2.5
5	3.5	4
6	2.5	3
7	2	3.5
8	1.5	3
9	1.5	2.5
10	2	2.5
11	3	2.5
12	2	2.5

21.8.4: Question 4

In a study of healthy subjects grouped by age (Younger: 19–50 years, Seniors: 65–75 years, and Longeval: 85–102 years), a researcher measured their vitamin B-12 levels (ng/L). All elderly subjects were living at home and able to carry out normal day-to-day activities. The following table shows vitamin B-12 levels for subjects in the young group, seniors, and longeval group. Assumptions are violated.

May we conclude, on the basis of these data, that the populations represented by these samples differ with respect to vitamin B-12 levels? Let α=0.01.

Table 21.54: Vitamin B-12 levels for subjects in the different age groups.

Young (19-50 years)	Senior (65-75 years)	Longeval (85-102 years)
230	319	
477	190	148
561	461	1941
347	163	128
566	377	145
260	190	174
247	335	495
241	371	161
442	460	149
491	440	409
279	520	229
334	256	183
300	238	400
230	137	348
215	452	175
260	437	262
349	411	
536	268	

All the knowledge belongs to God.

21.8.5: Question 5

A physical therapist conducted a study to compare three models of low-volt electrical stimulators. Nine other physical therapists were asked to rank the stimulators in order of preference. A rank of 1 indicates first preference. The results are shown in following table. We wish to know if we can conclude that the models are not preferred equally.

Table 21.55: Physical therapists' rankings of low-volt electrical stimulators.

Therapist	Model A	B	C
1	2	3	1
2	2	3	1
3	2	3	1
4	1	3	2
5	3	2	1
6	1	2	3
7	2	3	1
8	1	3	2
9	1	3	2

21.8.6: Question 6

Experiment was conducted to determine how the surface (grass, cement or rubberized running track) affected the time to sprint 40 yards. Twelve subjects were recruited for the study, and in order to compare the surface effect within each subject, all subjects ran on all three surfaces. To adjust for the lingering exhaustion effect of each run, following design was used. Two subjects were randomized to each sequence group in each square. The data resulting from this experiment is the time in seconds for each subject to sprint 40 yards and is shown in following table. Treatment level 1 represents cement, 2 as rubberized track and 3 as grass. Analyze the findings.

Table 21.56: William's design for three treatments

	Square Period I			Period II		
Group	1	2	3	1	2	3
1	1	2	3	3	2	1
2	2	3	1	1	3	2
3	3	1	2	2	1	3

21.8.7: Question 7

An investigator investigate how the breathing rates of cane toads (*Bufomarinus*) respond to conditions of hypoxia. Toads, the subjects, show two different kinds of breathing patterns, lung or buccal. The second factor was O_2 concentration, which had eight levels (0, 5, 10, 15, 20, 30, 40 and 50%). The response variable was the frequency of buccal breathing. Recognize the design along with Hypothesis Model.

21.8.8: Question 8

A cardiologist always interested to know the time between two hearts attacks along with client behavior about preventions. Probability of survival at 2nd heart attack is very low that's why doctor advised strict life schedule to the survival of first heart attack. These preventions hope so to increase the time of 2nd attack and save the life of patients. Time between two hearts attacks and

All the knowledge belongs to God.

patient number are given below (Table 21.57). Even patient number is allotted to female and odd for male. First six patients are those who takes low fat, next six takes medium and last six high. All males are involved in exercise, but female thinks they did not need any exercise/walk. All patient's data is collected with same city. Are preventions really helps to increase the time of second heart attack?

Table 21.57: Time of second heart attack.

Patient No	1	2	3	4	5	6	7	8	9
Time	12Y	9.7Y	10.2Y	7Y	17Y	9.6Y	8Y	5.9Y	6.11Y*
Patient No	10	11	12	13	14	15	16	17	18
Time	4.3Y	5.4Y	5.1Y	2Y	3Y	9M	1.11Y	7M	11M

*6.11Y=6 years and 11 months

INDEX

All the knowledge belongs to God.